

**UNIVERSITE MONTPELLIER II – Sciences et Techniques du Languedoc**

**THESE**

Pour obtenir le grade de

**Docteur de l'Université Montpellier II**

Discipline	Evolution, Ecologie, Ressources Génétiques, Paléontologie
Ecole doctorale	Systèmes Intégrés en Biologie, Agronomie, Géosciences, Hydrosciences et Environnement (SIBAGHE)

Préparée au sein de l'UMR Biologie et Génétique des Interactions Plantes-Parasites

Présentée et soutenue publiquement par

**Stéphanie ROBERT**

Le 3 avril 2012

**Emergence mondiale de la maladie des raies noires du bananier:  
histoire de l'invasion et stratégie de vie du champignon  
phytopathogène *Mycosphaerella fijiensis*.**

***Jury***

**Dr Thomas GUILLEMAUD, INRA, Sophia Antipolis - *Rapporteur***

**Dr Cécile ROBIN, INRA, Bordeaux - *Rapporteur***

**Dr Ruth HUFBAUER, Colorado State University, Fort Collins - *Examineur***

**Dr Yannis MICHALAKIS, IRD, Montpellier - *Examineur***

**Dr Jean CARLIER, CIRAD, Montpellier - *Directeur de thèse***

**Dr Virginie RAVIGNE, CIRAD, Montpellier - *Co-encadrante de thèse***

**Dr Catherine ABADIE (invitée), CIRAD, Montpellier - *Co-encadrante de thèse***

## **Résumé**

La compréhension des mécanismes écologiques et évolutifs sous-tendant l'émergence des maladies de plantes cultivées est un enjeu majeur, ancré dans la problématique générale des invasions biologiques. Nous l'abordons ici via l'étude du champignon phytopathogène *Mycosphaerella fijiensis*, responsable de la récente pandémie de maladie des raies noires du bananier. Une étude de la structure génétique mondiale à différents marqueurs neutres confirme l'aire d'origine sud-est asiatique de *M. fijiensis*, et suggère des scénarios d'invasion contrastés entre continents. L'invasion africaine découlerait ainsi d'un événement de fondation unique tandis que l'invasion de l'Amérique dériverait d'un mélange de différentes sources. Utilisant des méthodes d'analyse récentes (ABC « Approximate Bayesian Computation »), nous vérifions la pertinence de ces deux scénarios dans un cadre statistique et estimons des paramètres démographiques, qui suggèrent que la dissémination continentale s'est accompagnée de goulots d'étranglement drastiques. Par ailleurs, nous explorons les fondements du succès compétitif de *M. fijiensis*, qui a systématiquement supplanté l'espèce proche *M. musicola* précédemment établie, à l'aide d'une comparaison expérimentale des stratégies de vie des deux espèces. Notre étude suggère une importance cruciale et sous-estimée de la reproduction asexuée dans le succès de *M. fijiensis* et soulève des questions fondamentales dans le cadre de la théorie de l'évolution des histoires de vie. Nos travaux illustrent enfin l'intérêt d'une approche multidisciplinaire à visée fondamentale pour l'optimisation des stratégies de prévention et de lutte.

## **Mots clés**

Maladie émergente, invasion, génétique des populations, traits d'histoire de vie, *Mycosphaerella fijiensis*, champignon phytopathogène, bananier

## **Summary**

Deciphering ecological and evolutionary mechanisms underlying emerging diseases of crop plants is of major interest among current issues of biological invasions. Here, this issue is tackled through the study of the fungal plant pathogen *Mycosphaerella fijiensis*, responsible for the recent pandemic of black leaf streak disease of banana. A study of global genetic structure at different neutral markers confirms the South East Asian origin of *M. fijiensis* and reveals contrasted invasion scenarios between continents. In particular a single founder event underlies African invasion while American invasion would derive from an admixture between different sources. We used recent ABC (Approximate Bayesian Computation) methods to statistically validate these scenarios and estimate demographical parameters, that suggest the occurrence of drastic bottlenecks accompanying continental spread. Moreover we investigate the determinants of *M. fijiensis* competitive success, over the related species *M. musicola* previously established and systematically excluded by *M. fijiensis*. The experimental comparison of both species life history traits suggests an importance role for asexual reproduction in *M. fijiensis* establishment success, and raises fundamental issues about life history theory. Finally our work illustrates how such multidisciplinary approach with fundamental aims can provide key information for the optimization of prevention or management strategies.

## **Key words**

Emerging disease, biological invasion, population genetics, life history traits, *Mycosphaerella fijiensis*, fungal plant pathogen, banana

## *Remerciements*

J'ai eu la chance de mener cette thèse avec un encadrement de grande qualité, au sein d'une équipe et d'une unité dynamiques et conviviales, et je voudrais adresser un grand merci à toutes les personnes qui ont contribué à cette atmosphère idéale pour un thésard...

Merci...

... ou plutôt mille mercis à Jean, à Virginie et à Catherine pour votre remarquable appui et encadrement scientifique, au sein duquel je me suis sentie à l'aise pour progresser. Cela va de paire avec vos qualités humaines d'écoute et de sympathie et je vous remercie pour tout ça. Merci Jean entre autre pour ton souci et ta grande disponibilité, malgré ton rôle de chef d'équipe, Virginie ton super pouvoir à rendre limpide n'importe quelle notion ou théorie d'évolution compliquée en quelques mots simples, Catherine ton suivi et ton souci malgré la distance... ce ne sont que des exemples parmi tout le reste de ce que je retiendrai, qui nécessiterait de bien plus longs paragraphes...

Merci de m'avoir autant fait confiance et laissé assez libre cours à mon cheminement (du moins j'en ai l'impression, c'est sans doute le principal ;) ), et tout à la fois autant guidée et ré-aiguillée sur le bon chemin, celui qui va plus efficacement et rapidement à l'essentiel que l'autre, (le tout-embrouillé qui fait des détours inutiles dans toutes les directions et que j'ai souvent emprunté !)

A Marie, mille mercis pour la formidable acolyte que tu as été au laboratoire ! merci pour m'avoir coaché avec autant de patience et transmis un peu de ton précieux savoir-faire, ainsi qu'un appui sans faille, en matière de manipulations et triturations de tous ordres de notre champignon favori. Cet apprentissage aurait été tellement plus compliqué et tellement moins sympathique si tu n'avais pas été là...

A Françoise également pour ta sympathie, et ton aide précieuse au laboratoire (sans compter les carrés chocolat du goûter ;).

Merci à vous tous qui m'avez fourni un appui tout autant scientifique que sympathique : à Véro, Luc, Rémy, Anny, Steeve... pour leur appui technique au laboratoire, mais aussi à tous ceux à qui j'ai consommé du temps pour des questions et échanges scientifiques, François, Adri, Dounia, Enrique, Babeth, Benoît B....et j'oublie certainement du monde...

A mon comité de thèse Arnaud, Benoît F, Renaud, Johan pour avoir consacré beaucoup de votre temps à me fournir de précieux conseils et points de vue sur l'avancement des travaux.

Aux personnes rencontrées aux réunions du reid champignon et de l'ANR Emerfundis pour leurs retours et commentaires avisés sur mon travail.

Merci finalement à Cécile Robin et Thomas Guillemaud pour avoir accepté de consacrer du temps à être rapporteurs de ce travail, et de même à tous les membres du Jury qui ont accepté de l'évaluer.

Puisqu'une bonne partie de mon emploi du temps de thèse s'est déroulé dans l'environnement de la fac, je remercie toutes les personnes rencontrées là-bas qui ont contribué à faire de ces enseignements une expérience à part, vraiment agréable et enrichissante pour la thèse et pour la suite: à Agnès qui prend si bien soin de ses petits moniteurs, aux techniciens des bâtiments de TP, aux autres moniteurs...

Merci à Jean, Virginie et Catherine pour m'avoir permis de suivre une telle expérience.

Quant à l'environnement au sein de l'unité... j'adresse un grand merci à vous tous de BGPI avec qui j'ai passé de très bons moments !

A toute l'équipe 5, bananes, riz, cacao et compagnie, ce fut un plaisir pour moi de travailler au sein de cette équipe, merci pour tous les échanges que l'on a pu avoir.

Merci tout particulier à Geneviève pour ta sympathie et ta disponibilité en toutes circonstances ;)

A mes super colloqs de bureau pour l'atmosphère qu'on y a créé: Dounia bien sûr qui aime tellement ce bureau que tu as décidé de ne plus le quitter ;) , et ceux qui sont passés de façon plus éphémère mais toujours en nous laissant quelque chose de particulier : merci à Stelly, Antoine, Pinar, Mathilde, Ludovic... et bien sûr Enrique qui a débarqué à la fin avec ta bonne humeur, ton immense serviabilité, tes blagues ...disons... inimitables ! tu vas devenir le nouveau gardien du bureau, on compte sur toi !

A tous les copains thésards de BGPI avec qui on sait décompresser à la cantines et pendant les pauses du labo, et aussi pendant les plus grosses pauses hors du labo ;) ... aux deux docs Doudou (Juju et ta remarquable capacité à vous redonner le sourire même quand rien ne va plus... allezzz, intégre notre bureau), merci à Audrey pour ton amitié, à Pierre-O, Stella, Enrique, Virginie, Adri, Aurélie, Xavier, Flo, Mélanie, Imène, Jean-Phi... Aux copains des autres labos aussi Sarah, Elsa, Marie R... A ceux de la bière party du vendredi soir, et merci tout particulier à Fab pour ton amitié et ton aide phylogénétique !

Merci aux amis de Montpellier avec qui on se sent presque en famille quand on est loin de sa famille... à Marie et Fred pour m'avoir écoutée et toujours poussée vers l'avant même dans les moments où plus rien ne va...à toutes les super personnes de la bande Thibaut, ThomasG, Laurène, Marie-P, Eric, ThomasL, Guillelme, Xavier, Delphine... merci à vous tous pour toutes ces escapades les week-ends, ces bons repas et ces home cinémas, si agréables et si indispensables pour se ressourcer entre les semaines de travail et repartir du bon pied le lundi !

A mes amies de longue date Marine (merci de n'être ni en thèse ni dans un labo, et de me sortir complètement de cet univers de temps en temps !) et Anouck (merci d'être en thèse dans un labo et me faire mesurer ma chance quand tu me racontes tes aventures hallucinantes !)

Je dois enfin une immense reconnaissance à ma famille, trop dure à transcrire avec des mots ici, pour leur confiance et leur soutien sans faille depuis toujours... Pensée profonde pour ceux qui sont partis beaucoup beaucoup trop tôt...

Parce-que vous êtes un peu partout dans ces lignes, je remercie de tout coeur tous ceux qui m'ont soutenue par des gestes, des mots ou des pensées, et m'ont donné le courage de rebondir et d'aller au bout de ce projet de thèse, dans un moment où la vie vient nous rappeler cruellement qu'elle peut basculer à tout moment et pour n'importe qui... Merci également à Jean, à Philippe et toutes les personnes du Cirad qui se sont coupées en quatre pour prolonger ma thèse, merci de m'avoir accordé votre confiance pour la terminer dans d'aussi bonnes conditions qu'elle a commencé.

# SOMMAIRE

<b>INTRODUCTION .....</b>	<b>9</b>
<b><u>I. Contexte : au carrefour entre maladies des plantes cultivées et invasions biologiques .....</u></b>	<b>11</b>
<u>I.1. Les maladies émergentes des cultures : défis et enjeux d'étude .....</u>	11
<u>I.2. Des invasions biologiques en agriculture .....</u>	15
<b><u>II. Comprendre le succès écologique et évolutif des maladies émergentes fongiques.....</u></b>	<b>19</b>
<u>II.1. Quel cadre conceptuel commun ? .....</u>	19
<u>II.2. Retracer l'histoire des populations envahissantes .....</u>	29
II.2.1. Perspectives directement appliquées .....	29
II.2.2. L'importance des méthodes employées .....	31
II.2.3. Les questions supplémentaires posées par les maladies émergentes fongiques de cultures .....	37
II.2.4. Quels scénarios d'invasion caractérisent les pathogènes fongiques émergents?.....	41
<u>II.3. Etablissement et expansion des populations : déterminer les caractères clés de l'interaction entre le pathogène et l'environnement .....</u>	43
II.3.1. Importance des traits d'histoire de vie liés à la colonisation et à la compétition lors des invasions .....	45
II.3.2. Réinvestir les méthodes théoriques et expérimentales de l'épidémiologie .....	47
<u>II.4. Vers l'intégration de l'histoire démographique des populations et de leurs stratégies de vie pour décrypter le succès d'invasion .....</u>	51
<b><u>III. La maladie des raies noires du bananier causée par <i>Mycosphaerella fijiensis</i> : un modèle de maladie émergente fongique à l'échelle mondiale .....</u></b>	<b>53</b>
<u>III.1. Impacts en agriculture .....</u>	53
<u>III.2. Suivi historique de l'émergence mondiale .....</u>	55



III.3. Biologie du pathosystème bananier/ <i>Mycosphaerella fijiensis</i> .....	57
III.4. Structure et diversité génétique des populations .....	61
<b>IV. Objectifs de la thèse</b> .....	<b>63</b>
 <b>CHAPITRE 1. Reconstitution de la propagation mondiale de <i>M. fijiensis</i></b> .....	<b>67</b>
<u>Article 1.</u> Contrasting introduction scenarios among continents in the worldwide invasion of the banana fungal pathogen <i>Mycosphaerella fijiensis</i> .....	71
<u>Article 2.</u> Optimized genotyping with microsatellite markers in the fungal banana pathogen <i>Mycosphaerella fijiensis</i> ( <i>Mycosphaerellaceae</i> ). .....	99
 <b>CHAPITRE 2. Histoire d'invasion en Afrique et en Amérique: tests des scénarios d'introduction et caractérisation de la propagation intracontinentale</b> .....	<b>103</b>
<u>Article 3.</u> Single introduction followed by multiple drastic bottlenecks characterize the emergence and spread of a banana wind-born pathogen across Africa .....	107
<u>Etude préliminaire.</u> Histoire de l'invasion de <i>M. fijiensis</i> en Amérique : processus d'introduction et d'établissement en contextes continental et insulaire. ....	145
 <b>CHAPITRE 3. Comparaison des stratégies de vie des cercosporioses du bananier : un rôle dans le remplacement par <i>Mycosphaerella fijiensis</i> de <i>M. musicola</i> ?</b> .....	<b>161</b>
<u>Article 4.</u> Asexual life strategies differ between two related fungal pathogens in competition on the same plant. The case of <i>Mycosphaerella fijiensis</i> and <i>M. musicola</i> on banana .....	167
 <b>DISCUSSION GENERALE et PERSPECTIVES</b> .....	<b>207</b>
 <b>REFERENCES</b> .....	<b>229</b>





# **INTRODUCTION GENERALE**



## **I. Contexte : au carrefour entre maladies des plantes cultivées et invasions biologiques**

### **I.1. Les maladies émergentes des cultures : défis et enjeux d'étude**

La consommation alimentaire mondiale est extrêmement dépendante des quatre principales denrées de base que sont le blé, le riz, le maïs et la pomme de terre, ainsi que d'autres produits comme la banane, le café et le cacao qui représentent des sources d'échanges et de revenus majeures pour les pays producteurs (Anderson *et al.* 2004). C'est pourquoi les maladies infectieuses causées par des virus, des champignons ou des bactéries, susceptibles d'affecter ces productions sont au cœur d'importantes recherches appliquées et théoriques (Anderson *et al.* 2004, Gilligan 2008, Stukenbrock & McDonald 2008). C'est d'autant plus vrai durant ces dernières décennies où l'émergence de maladies des cultures de plus en plus destructrices ne cesse de s'accélérer (Gilligan *et al.* 2008). Citons par exemple les récentes épidémies de mosaïque Africaine du manioc en Afrique de l'Ouest (Legg 1999), le chancre bactérien des agrumes en Floride (Gottwald *et al.* 2001) ou encore les récentes épidémies particulièrement sévères de la souche UG99 de rouille noire du blé qui se répandent en Afrique et menacent d'autres continents (Singh *et al.* 2011).

Selon la définition proposée par Anderson *et al.* (2004), on appelle « maladie émergente » toute maladie causée par un agent pathogène (i) dont l'incidence a augmenté par expansion de son aire de répartition ou élargissement de sa gamme d'hôtes, (ii) dont la pathogénicité a changé, (iii) dont l'espèce vient d'évoluer ou encore (iii) que l'on vient de découvrir ou de reconnaître, ces différentes possibilités pouvant naturellement se combiner. Le caractère nouveau, voire soudain, du phénomène est commun à tous les cas d'émergence de maladies, tout comme les importants dégâts qu'elles peuvent causer. On observe de plus qu'une majeure partie des maladies émergentes implique des introductions de pathogènes dans de nouvelles régions géographiques selon le premier cas défini par Anderson *et al.* (2004).

Les champignons et organismes assimilés (oomycètes) sont, après les virus, les plus grands responsables de maladies émergentes de plantes. Toutes les cultures ou presque sont affectées par une ou plusieurs espèces de champignons phytopathogènes, des céréales aux arbres fruitiers (parmi lesquels les caféiers et cacaoyers), jusqu'aux plantes et arbres



ornementaux. Toutes les parties des plantes peuvent être affectées, y compris fruits et graines, par une grande variété de symptômes (lésions, nécroses, chancres...). Leur impact est donc colossal en agriculture sur tous les continents, où ils peuvent réduire sévèrement les rendements et la qualité des produits agricoles, voire condamner toute production (Anderson *et al.* 2004). Les implications humaines peuvent être dramatiques comme le montre l'exemple européen bien connu de la grande famine du XIX<sup>ème</sup> siècle en Irlande, générée par l'oomycète *Phytophthora infestans* qui a anéanti brusquement les cultures de pommes de terre à la base de l'alimentation (Birch & Whisson 2001).

Les activités anthropiques ont une grande responsabilité dans la recrudescence actuelle des maladies émergentes de tout ordre, notamment via l'augmentation des échanges (Anderson *et al.* 2004, Desprez-Loustau *et al.* 2009). En particulier, le commerce international des produits agricoles, a permis aux champignons susceptibles de les infecter d'être disséminés à l'échelle mondiale. De plus la globalisation de l'agriculture s'est manifestée par des modifications profondes des pratiques agricoles durant les dernières décennies. L'uniformisation et l'intensification des agrosystèmes à l'échelle mondiale ont joué un rôle particulièrement favorable pour l'émergence de pandémies (Stukenbrock & McDonald 2008). En réponse, diverses stratégies de prévention et de lutte se sont alors développées, comme les systèmes phytosanitaires de quarantaine destinés à prévenir de nouvelles introductions de matériel végétal infecté. Une variété de pratiques culturales est de plus employée dans les agro-écosystèmes modernes afin de réduire la sévérité des maladies déjà installées, par exemple les pulvérisations de pesticides, la création de variétés hôtes résistantes ou la rotation des cultures (Anderson *et al.* 2004).

Mais alors que le risque de nouvelles émergences fongiques ne cesse de croître (Grünwald & Goss 2011), le développement de méthodes de contrôle vraiment durable est encore souvent à l'étude, et une meilleure compréhension théorique des facteurs ayant déclenché et favorisé les émergences de ces champignons s'avère pour cela nécessaire. Cette compréhension passe aussi bien par les caractéristiques biologiques de l'agent pathogène lui-même que par celles de l'environnement et par l'impact des activités humaines (Desprez-Loustau *et al.* 2009). L'ensemble de ces considérations n'est pas sans rappeler celles soulevées par les invasions biologiques d'une manière générale (Lodge 1993). Pourtant l'idée d'étudier ces deux phénomènes de façon interconnectée, proposée et mise en valeur par Desprez-Loustau *et al.* (2007), est relativement récente. Ce parallèle mérite d'être explicité ici car nous l'exploiterons tout au long de ce travail.



## **I.2. Des invasions biologiques en agriculture**

Par invasion biologique, nous entendons ici tout établissement et expansion réussie d'une espèce dans une région géographique où elle n'était pas indigène (Facon *et al.* 2006). Suivant cette définition, de nombreux agents pathogènes des cultures se révèlent finalement être des envahisseurs récents, disséminés au-delà de leurs anciennes limites géographiques (Desprez-Loustau *et al.* 2007). Les études d'invasions biologiques, notamment de plantes ou d'animaux, sont en croissance exponentielle depuis les dernières décennies (Pysek *et al.* 2010). Bon nombre de ces études s'attachent à éclaircir les mécanismes et facteurs écologiques et/ou évolutifs responsables du succès d'invasion (Gurevitch *et al.* 2011, Handley *et al.* 2011). Il est intéressant de relever d'un point de vue historique que les invasions ont d'abord été appréhendées sous un angle purement écologique. L'intégration d'aspects évolutifs reste relativement récente, en dépit des indices évidents de changements évolutifs rapides dans la morphologie, le comportement ou les traits d'histoire de vie de certaines espèces relevés depuis longtemps (Reznick & Ghalambor 2001, Lambrinos 2004). Au final, il est à l'heure actuelle bien établi que les invasions biologiques fournissent des opportunités inégalées d'étudier les processus évolutifs à l'œuvre sur les échelles de temps contemporaines, dites « écologiques » (Reznick & Ghalambor 2001, Lee 2002, Lambrinos 2004, Facon *et al.* 2006, Handley *et al.* 2011).

Or, dans les agro-écosystèmes, peuvent s'opérer des changements évolutifs tout aussi – voire plus – rapides et drastiques que dans les écosystèmes naturels, précisément en raison de l'impact prépondérant des activités humaines sur la structure et composition des agro-écosystèmes. La création de nouvelles variétés hôtes ou les modifications de pratiques agricoles exercent de nouvelles pressions de sélection pour les organismes pathogènes, favorisant les adaptations rapides (Stukenbrock & McDonald 2008). C'est ainsi qu'en plus des enjeux appliqués qu'elles posent, les émergences de maladies apparaissent tout autant comme des systèmes modèles pour l'étude de phénomènes d'évolution rapide. Au même titre que d'autres organismes, les agents pathogènes de plante peuvent permettre de tester des hypothèses plus ou moins propres aux mécanismes d'invasions ou générales en biologie évolutive.

Il apparaît donc que l'étude des émergences de maladies de plantes cultivées soulève des problématiques à la fois fondamentales et appliquées semblables aux invasions biologiques, et fera souvent appel à des démarches et outils d'études similaires (Desprez-Loustau *et al.* 2007). L'étude des mécanismes d'invasions biologiques en général peut





apporter beaucoup à celle des maladies émergentes. Ainsi un nombre croissant d'études récentes (revues dans Gurevitch *et al.* 2011) oriente leurs efforts vers des tentatives d'unification qui permettraient d'aborder tous les cas d'invasions biologiques selon les mêmes cadres conceptuels, malgré leur grande variété et leurs spécificités taxonomiques et environnementales. L'idée est que ces cadres théoriques puissent servir de base commune de réflexion sur les déterminants communs des invasions. L'étude des émergences de maladies doit naturellement pouvoir en bénéficier et les réinvestir à son profit. De manière réciproque, l'étude de cas d'envahisseurs pathogènes en agriculture peut apporter beaucoup à la compréhension des invasions en général, et notamment à ces tentatives d'unification conceptuelle. Les invasions d'agro-écosystèmes par des agents pathogènes doit constituer, au même titre que d'autres organismes peut-être plus charismatiques, des systèmes modèles dont la compréhension peut contribuer à l'élaboration de tels cadres théoriques généraux.

De telles considérations nous paraissent particulièrement vraies dans le cas précis des champignons pathogènes, encore beaucoup trop peu abordés dans la littérature théorique des invasions biologiques (Pysek *et al.* 2008). Ceci s'explique notamment par le cloisonnement des disciplines qui s'observait jusqu'alors. Les pathogènes de plantes cultivées représentaient plutôt l'objet d'étude de la phytopathologie, discipline qui, à ses débuts, cherchait prioritairement à recenser et décrire les nouvelles maladies pour mettre au point des méthodes de lutte. Mais une autre raison majeure serait surtout notre manque de connaissance quant à la biologie précise de ces agents pathogènes. Les champignons sont parmi les organismes les plus diversifiés d'un point de vue écologique et évolutif (Desprez-loustau *et al.* 2007). Une connaissance approfondie du cycle de vie et des particularités biologiques d'un organisme envahissant est bien sûr un pré-requis pour l'étude de son succès d'invasion. Or les champignons présentent des cycles de vie souvent complexes, ayant chacun leurs spécificités. Plusieurs plantes hôtes peuvent être impliquées, ainsi que l'alternance entre différents modes de reproduction et différentes structures de dispersion ou de survie associées. On peut citer à titre d'exemple le cycle de la rouille du peuplier causée par le champignon *Melampsora larici populina*, alternant la production asexuée ou sexuée de 5 types de spores différents, sur peuplier (hôte principal) et mélèze (Pei *et al.* 2005).

Or cette incroyable diversité de cycles et stratégies de vie qui caractérise déjà les agents pathogènes en général, offre pourtant des opportunités inégalées d'explorer des questions fondamentales de biologie évolutive, telles que celles liées à l'évolution des modes de reproduction sexuée ou asexuée ou encore à l'évolution des traits d'histoire de vie clés liés



au rythme de la reproduction et à la dispersion. De par les millions de spores que certaines espèces sont en mesure de produire à partir d'une seule plante infectée, et la rapidité de leurs cycles (souvent plusieurs générations par an), on pourrait s'attendre à ce que ces organismes puissent s'adapter de façon particulièrement rapide. Le potentiel d'études fondamentales qu'ils offrent dans ce contexte d'invasions récentes, reste encore à explorer.

## **II. Comprendre le succès écologique et évolutif des maladies émergentes fongiques**

Lorsqu'on se retrouve confronté à une nouvelle maladie émergente fongique, causée par un pathogène envahissant parfois récemment découvert et peu décrit, comment peut-on élucider les facteurs écologiques et évolutifs ayant permis et favorisé cette émergence ? Nous commencerons par nous resituer dans un cadre conceptuel éco-évolutif général, appréhendé sous l'angle spécifique des invasions de pathogènes fongiques en agriculture, afin de mettre en lumière les problématiques centrales qui peuvent guider nos recherches. Partant de là, nous verrons quelles méthodes peuvent faire leurs preuves pour répondre à ces questions. Nous évoquerons les apports possibles d'un point de vue immédiatement appliqué pour optimiser les stratégies de contrôle du pathogène considéré, et d'un point de vue plus théorique pour comprendre les mécanismes d'invasion en général.

### **II.1. Quel cadre conceptuel commun ?**

Une question centrale en biologie des invasions, lorsque l'on cherche à synthétiser les avancées effectuées par l'ensemble des études, serait donc « pourquoi certaines introductions réussissent et d'autres non » ? En d'autres termes, peut-on mettre en évidence des déterminants généraux qui permettraient de prédire la réussite des invasions biologiques ?

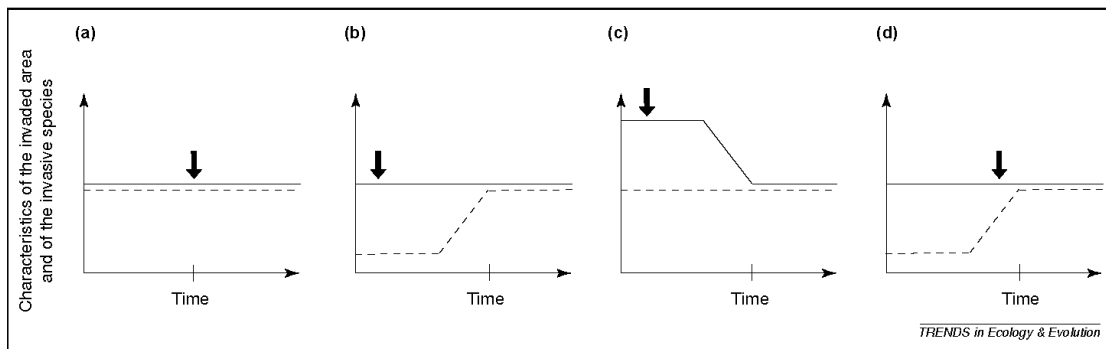
L'identification de traits biologiques typiques des envahisseurs les plus performants d'une part (qui seraient liés par exemple aux capacités de dispersion, capacité à faire de la reproduction asexuée, tolérance à l'hétérogénéité de l'environnement), et des caractéristiques générales des environnements envahis d'autre part, constitue depuis longtemps un défi auquel de nombreuses études ont tenté d'apporter des réponses (par ex., Kolar & Lodge 2001, Richardson & Pysek 2006, van Kleunen *et al.* 2010). De nombreuses études s'attachent ainsi à comparer introductions réussies et échecs, ou encore espèces envahissantes et non-



envahissantes (revues dans van Kleunen *et al.* 2010). Dans le domaine des champignons phytopathogènes envahissants, il a pu être montré que des traits liés à la dispersion à longue distance et à la reproduction sexuée pouvaient être de bons prédicteurs du succès d'invasion chez les espèces forestières (Philibert *et al.* 2011). Mais si ces efforts ont permis de progresser notablement dans la connaissance des invasions chez tel ou tel organisme, ils ont rarement permis d'extraire une liste de déterminants communs à tous les envahisseurs ou à tous les environnements envahis (même s'il a toutefois été suggéré que les perturbations du milieu pourraient être une caractéristique fréquente des environnements envahis (Gurevitch *et al.* 2011).

Une façon plus générale d'aborder différents cas d'invasions serait donc de ne pas considérer séparément l'envahisseur et son environnement, mais bien l'interaction entre eux. La réussite d'une invasion biologique implique en effet l'existence d'une affinité entre un organisme et l'environnement dans lequel il est introduit, et c'est cette affinité qu'il convient de caractériser et d'expliquer, comme l'ont mis en valeur plusieurs auteurs proposant des cadres théoriques à l'étude des invasions (Facon *et al.* 2006, Barney & Whitlow 2008, Gurevitch *et al.* 2011). C'est aussi ce que suggérait déjà, par une approche purement écologique, le concept d'opportunité de niche de Sea & Chesson (2002) qui soulignent l'importance d'une adéquation entre la niche d'un envahisseur et les opportunités offertes par son nouvel environnement, particulièrement les facteurs ressources, ennemis naturels et environnement physique. Sea & Chesson (2002) ont alors suggéré d'aborder l'écologie des invasions comme partie intégrante des concepts de l'écologie des communautés. Plus largement, nous pensons comme Gurevitch *et al.* (2011) que l'étude des invasions doit se placer dans un cadre incluant non seulement l'écologie des populations et des communautés, mais également de manière cruciale des concepts évolutifs, considérant à la fois l'évolution des populations envahissantes et des populations résidentes. Ceci souligne également que l'on doit considérer la dynamique des populations et leur variabilité au sein des espèces, et non plus les caractéristiques générales de l'espèce dite « envahissante ».

Facon *et al.* (2006) ont ainsi proposé les premiers d'aborder le processus d'invasion comme la réalisation d'une adéquation efficace entre l'envahisseur et le nouvel environnement. Si cette affinité apparaît bien sûr propre à chaque système, elle se base sur des processus écologiques et évolutifs basiques qu'il faut alors élucider. Facon *et al.* (2006) proposent ainsi différents scénarios de base qui décomposent ces mécanismes selon des schémas simples, lesquels sont susceptibles de se combiner pour expliquer de façon réaliste la réussite de tel ou tel cas d'invasion (Figure 1).



**Figure 1: Scénarios d’invasions théoriques proposés par Facon et al. (2006).**

Les caractéristiques de l’environnement envahi et de l’envahisseur sont représentées respectivement par une ligne pointillée et un trait plein, qui se superposent lorsque une adéquation entre les deux est réalisée. Une flèche épaisse indique le moment du premier contact potentiel entre environnement envahi et envahisseur (par migration), tandis qu’un tiret vertical indique le début du processus d’invasion par établissement de l’envahisseur. **(a) Scénario 1** : seul un changement du régime de migration suffit à initier l’invasion; **(b) Scénario 2** : un changement dans les caractéristiques de l’environnement envahi suffira à initier l’invasion; **(c) Scénario 3** : un changement évolutif chez l’envahisseur permet d’initier l’invasion; **(d) Scénario 1+2** : exemple de scénarios combinés, où l’invasion est initiée suite à un changement du régime de migration plus une modification des caractéristiques de l’environnement. (Source: Facon et al. 2006)

- ❖ Dans le premier scénario possible, le seul changement du régime de migration peut initier l'invasion en permettant le contact entre envahisseur et nouvel environnement. Si l'envahisseur était jusque-là absent de cette région géographique, c'est uniquement parce qu'il n'avait pas eu les moyens d'y parvenir. Un nouveau franchissement de barrière aurait pu alors s'effectuer notamment via les mouvements humains à grandes distances. L'établissement et l'expansion efficace de l'envahisseur à son arrivée supposent alors l'adaptation préalable de l'envahisseur à ce nouvel environnement, proche de son environnement d'origine, et aucune modification génétique ni écologique n'est nécessaire.
- ❖ Le deuxième scénario suppose que malgré la migration possible, voire régulière, de l'envahisseur jusqu'à cette nouvelle zone, l'affinité entre l'envahisseur et l'environnement dans lequel il arrivait n'était pas réalisée, prévenant tout établissement. L'affinité s'est alors réalisée suite à des modifications des conditions environnementales, qu'il s'agisse de perturbation et restructuration totale de l'environnement par les activités anthropiques, ou de changements climatiques liés par exemple au réchauffement global.
- ❖ Enfin le troisième scénario suppose que c'est cette fois-ci l'envahisseur qui a évolué dès son arrivée dans le nouvel environnement afin d'être en mesure de s'y établir et se propager durablement. Les pressions de sélection du nouvel environnement qui ont contraint cette adaptation peuvent être de nature abiotique et/ou biotique. Dans le cas des agents pathogènes émergents, Stukenbrock et McDonald (2008) proposent notamment les mécanismes évolutifs les plus fréquents qui seraient à l'origine des espèces pathogènes dominant les agrosystèmes modernes, tel que la domestication du pathogène en même temps que celle de son hôte (« host tracking ») ou encore les sauts d'hôte, le transfert horizontal de gènes et l'hybridation. De tels mécanismes, notamment liés à la domestication des plantes cultivées, ont eu lieu sur des échelles de temps anciennes, mais peuvent aussi se produire plus rapidement de façon contemporaine dans le contexte de globalisation de l'agriculture et recrudescence des invasions de nombreux pathogènes sur les mêmes cultures (Stukenbrock & McDonald 2008). L'ampleur de telles évolutions chez les agents pathogènes est encore relativement méconnue (Arnold 2004), mais un exemple bien documenté chez un pathogène d'arbre, est l'hybridation entre des espèces du genre *Ophiostoma* sp. (*O. ulmi* et *O. novo-ulmi*) qui s'est produite de façon abrupte au cours du siècle dernier,





- ❖ générant de sévères pandémies de graphiose de l'orme en Europe et Amérique du Nord (Brasier 2001).

Les 3 scénarios proposés par Facon *et al.* (2006) peuvent naturellement se combiner pour rendre compte des cas d'invasion observés, et ce serait particulièrement vrai dans le cas des invasions en agriculture, où changements évolutifs et écologiques sont susceptibles d'intervenir, fortement conditionnés par l'Homme. En effet, après avoir domestiqué des plantes à partir du compartiment sauvage, et par la même occasion domestiqué des pathogènes, nous avons modifié les environnements de façon globale en y introduisant ces plantes domestiquées au sein d'agro-écosystèmes. Ceux-ci sont souvent caractérisés par une grande promiscuité des plantes hôtes, une grande uniformité de structure et composition, et notamment une grande homogénéité génétique, qui peuvent être susceptible d'atténuer l'inadéquation (« mismatch ») qu'on attend généralement lorsqu'une espèce est introduite dans une aire éloignée. De plus les mouvements massifs de matériel végétal engendrés, entraînant avec eux les pathogènes, peuvent permettre de réunir fréquemment le pathogène et sa plante hôte dans de nouvelles régions (Stukenbrock *et al.* 2008, Desprez-Loustau *et al.* 2009).

Ainsi, chez les pathogènes, les trois scénarios peuvent entrer en jeu de façon imbriquée, et avec l'intervention d'un niveau supplémentaire de complexité qu'il convient d'intégrer aux scénarios de Facon *et al.* (2006). La majorité des études s'intéressaient jusqu'ici aux processus se déroulant au sein de l'aire envahie suite à une introduction, une vision qui se justifie par l'idée que c'est à leur arrivée dans de nouvelles régions que les populations envahissantes vont être soumises à de nouvelles pressions de sélection. Cependant, on peut envisager que des changements évolutifs aient eu lieu avant l'introduction des populations envahissantes auxquelles on s'intéresse, donc avant l'étape de mise en contact (Lee & Gelembiuk 2008). Il émerge récemment de la littérature que des changements évolutifs et/ou écologiques préalables à la dissémination à grande échelle, voire au sein même de l'aire d'origine, pourraient rendre compte de la réussite de bon nombre d'invasions et émergences qui ont lieu dans des contextes d'environnements anthropisés généralement homogènes et uniformisés à échelle globale. C'est ce que propose par exemple le scénario « tête de pont », selon lequel une population envahissante particulièrement performante, la population « tête de pont », serait la source principale ayant relayé l'invasion dans d'autres aires envahies aux environnements semblables (Lombaert *et al.* 2010). Si la population « tête de pont » a subi un changement adaptatif, alors les populations envahissantes qui en dérivent pourraient être



préalablement adaptées à l'environnement dans lequel elles sont introduites. Ceci s'avère en effet plus parcimonieux d'un point de vue évolutif qu'un scénario décrivant plusieurs événements d'adaptations, à chaque introduction dans une nouvelle région géographique à partir de populations sources de l'aire d'origine. Un tel scénario, montré chez la coccinelle *Harmonia axyridis*, pourrait probablement rendre compte de nombreux cas d'invasions par des pestes des cultures (Lombaert *et al.* 2010, Estoup & Guillemaud 2010). Le scénario AIAI (« Anthropogenically Induced Adaptation to Invade », Hufbauer *et al.* 2011) met plus spécifiquement l'accent sur le rôle des modifications anthropiques contemporaines et des nouvelles pressions de sélection qu'elles engendrent au sein même de l'aire d'origine. Les populations natives qui s'adaptent alors à ces environnements anthropisés pourraient être fréquemment la source d'introductions vers d'autres régions par les activités humaines, et seraient alors préalablement adaptées aux environnements anthropisés dans lesquelles elles se retrouvent introduites, donc particulièrement aptes à envahir. Ces scénarios proposés récemment, pourraient rendre compte de bon nombre de cas d'invasion de pathogènes en agriculture puisqu'on a justement affaire à des milieux profondément modifiés par l'homme, et ce de façon homogène à l'échelle globale. Ces questions méritent donc d'être explorées et documentées par des études de cas.

Si des cas d'évolutions rapides ont souvent été décrits durant l'établissement et l'expansion d'envahisseurs (Whitney & Gabler 2008), il n'est pas si simple de démêler si les changements mis en évidence dans des traits phénotypiques et d'histoire de vie sont bien le reflet d'évolutions adaptatives ou si ils sont le résultat de processus évolutifs neutres. Même dans des cas bien démontrés d'adaptation, savoir si elle a réellement joué un rôle pour la réussite de l'invasion n'est pas forcément clair puisque cette adaptation peut aussi juste coïncider avec l'invasion ou en être une conséquence, sans forcément l'avoir favorisé (Handley *et al.* 2011). Le hasard peut en effet tenir une place très importante dans l'évolution des populations envahissantes. Le processus d'expansion d'aire de distribution est en effet particulier dans le cas des invasions en raison de l'intervention prédominante de l'Homme. Des scénarios de dissémination complexes sont souvent montrés, impliquant sauts à longue distance et successions de goulots d'étranglements et d'introductions multiples dans le temps et dans l'espace. L'étape de dissémination, de par sa nature souvent stochastique, peut donc faire intervenir de concert différentes forces évolutives : notamment une forte dérive due à l'échantillonnage aléatoire d'un petit nombre d'individus dans une population source (goulot d'étranglement), ou encore la migration en tant que force évolutive, dans des cas d'introductions multiples conduisant à des mélanges de populations différenciées (admixture).



Une étude intégrée de l'histoire du processus d'invasion, depuis les populations sources jusqu'aux populations envahissantes en passant par l'histoire de dissémination et d'introduction apparaît alors nécessaire pour comprendre les facteurs qui ont fait la réussite de bon nombre d'invasions (Estoup & Guillemaud 2010).

Lorsque l'on est confronté à un cas particulier d'émergence de pathogène, le problème qui se pose vise alors à comprendre dans quel(s) scénario(s) de Facon *et al.* (2006) se replace cet envahisseur. Comment distinguer la contribution relative des différents facteurs écologiques et forces évolutives, et le moment auquel ils interviennent ? Une démarche générale à adopter peut être de décrypter le processus d'invasion suivant ses différentes étapes :

- tout d'abord reconstruire l'histoire des populations envahissantes d'un point de vue géographique, génétique et démographique : sources, routes de dissémination et scénarios (nombre, lieux) d'introductions ;

- dans un deuxième temps comprendre le fonctionnement de l'adéquation réussie entre les populations envahissantes et l'agro-écosystème et en particulier identifier les traits du pathogène et les caractéristiques de l'environnement qui s'avèrent clés pour la réussite de l'établissement et de l'expansion ;

- l'objectif final est de pouvoir d'intégrer ces deux aspects afin de comprendre si une évolution adaptative de ces traits clé a été nécessaire ou non pour permettre la réussite de l'invasion.

Cette démarche est développée dans la suite étape par étape. On verra que les problématiques soulevées à chaque étape se posent de façon particulière et doivent faire face à des défis spécifiques dans le cas des pathogènes émergents en agriculture, en raison de notre manque fréquent de connaissances quant à la biologie de ces organismes.

## **II.2. Retracer l'histoire des populations envahissantes**

### **II.2.1. Perspectives directement appliquées**

D'un point de vue plus directement appliqué, la reconstruction de l'histoire des populations fournit des informations précieuses pour la mise en place immédiate des stratégies de contrôle dans les zones affectées et des stratégies de prévention dans les zones encore indemnes. Notamment, la connaissance des populations sources renseigne également



sur leur composition génétique et l'environnement biotique et abiotique auquel elles étaient *a priori* adaptées, par exemple les pesticides utilisés pour les contrôler. Dans le cas d'un scénario « tête de pont », connaître cette population envahissante « tête de pont » à partir de laquelle le pathogène se serait disséminé à large échelle, apparaît particulièrement important. De plus, en retraçant le nombre d'introductions qui a été nécessaire à initier l'invasion et leur origine, on peut en apprendre plus sur les modes de dissémination qui sous-tendent l'invasion, notamment s'il s'agit d'introductions via les activités humaines ou par des moyens de dispersion naturels. L'ensemble de ces informations sont précieuses pour estimer la pertinence des plans de surveillance et des mesures de quarantaine, et prédire les risques futurs (Estoup & Guillemaud 2010).

### **II.2.2. L'importance des méthodes employées**

#### **❖ Méthodes directes**

Au cours du processus d'expansion géographique, il n'est pas rare que l'on ait enregistré de façon plus ou moins précise des données relatant les premières détections et l'expansion d'une maladie dans de nouvelles régions. L'ensemble constitue ainsi un suivi spatio-temporel direct de l'expansion de la maladie, mais celui-ci n'est pas forcément exploitable en soi. L'arrivée d'une maladie peut être bien antérieure à sa première détection, surtout chez des organismes fongiques où la différenciation morphologique des espèces basée sur l'observation des symptômes n'est pas toujours évidente. Malgré tout, ces données peuvent servir de repères et permettre d'émettre des hypothèses, qui pourront être explorées plus précisément grâce à la mise en œuvre de méthodes indirectes basées sur des marqueurs génétiques. Elles pourront donc constituer de précieuses informations à intégrer avec les méthodes indirectes (Estoup & Guillemaud 2010).

#### **❖ Méthodes indirectes**

Les méthodes indirectes doivent être employées pour retracer l'histoire des populations envahissantes et les relier à leurs populations sources. Mais une population envahissante que l'on peut échantillonner dans l'aire d'invasion, aura donc souvent subi une suite d'évènements stochastiques au cours de son histoire : échantillonnage aléatoire d'une partie de la diversité des populations sources (effet de fondation), goulots d'étranglements répétés, admixture avec





d'autres populations. Tous ces processus passés vont laisser d'importantes signatures dans la diversité génétique neutre des populations, qui peuvent nous permettre de les retrouver par l'emploi des marqueurs adaptés, mais également brouiller les pistes qui nous permettraient de relier populations envahissantes à leurs populations sources par les méthodes descriptives évoquées (Keller & Taylor 2008, Estoup & Guillemaud 2010). Il peut donc être crucial de combiner et confronter autant que possible différentes méthodes basées sur des marqueurs à différentes résolutions (Sunnucks 2000, Jombart *et al.* 2010).

On fait premièrement appel aux méthodes de génétique des populations de plus en plus puissantes (Excoffier et Heckel 2006) qui permettent de reconstruire les routes d'invasion à partir de la distribution et structuration de la diversité génétique des populations actuelles de l'aire d'origine et d'invasion (Estoup & Guillemaud 2010). Les marqueurs microsatellites notamment, de part leur évolution rapide, peuvent enregistrer des processus populationnels sur les échelles de temps courtes caractérisant les invasions (Selkoe & Toonen 2006, Dutech *et al.* 2007). Ils apparaissent particulièrement adaptés pour explorer la diversité et la structure des populations envahissantes afin de les relier à leurs populations sources potentielles. De nombreuses méthodes descriptives sont ainsi basées ainsi sur les modifications de fréquences alléliques au niveau populationnel, caractérisées par le calcul d'indices de diversité intra-populations et/ ou de différenciations inter-populations (de Meeûs *et al.* 2007). La construction d'arbres de populations permet de visualiser des distances génétiques, calculées selon différents critères. Les méthodes un peu plus récentes de classification (« clustering ») et d'assignation, basées sur des modèles (telles que celle implémentée dans STRUCTURE, Pritchard *et al.* 2000) ou non (analyses multivariées, par ex. Jombart *et al.* 2010) peuvent s'avérer particulièrement puissantes à conditions que l'on dispose d'un bon échantillonnage des aires d'origine et envahie, et d'une résolution de marqueurs adaptée.

Par ailleurs, dans certains cas, les méthodes de reconstruction d'arbres basées sur des séquences peuvent s'avérer précieuses pour retracer l'histoire des invasions. Selon l'ancienneté des processus d'invasion, et le degré de polymorphisme des séquences utilisées sur les échelles spatiotemporelles considérées, les méthodes d'inférence issues de la phylogénie moléculaire, basées sur les arbres et réseaux phylogénétiques (Posada & Crandall 2001, Morrison 2005) peuvent apporter de précieuses informations sur les relations phylogénétiques des individus, et donc les voies de disséminations empruntées (Hare 2001, Avise 2009, Brito & Edwards 2009). L'idée est que les séquences, en général moins polymorphes que les microsatellites car à évolution moins rapide, sont moins homoplasiques



et donc pertinentes pour reconstruire des topologies (arbres ou réseaux). Elles peuvent donc également conserver plus longtemps la trace d'un polymorphisme partagé entre individus de l'aire envahie et de l'aire d'origine par exemple. Notons que, là encore, les champignons ont été longtemps sous-représentés dans le domaine de la phylogéographie (Beheregaray 2008), par manque total de connaissances sur la structure de leurs populations. Une idée commune était que les champignons ne présentent pas de populations structurées du fait de leurs grandes capacités de dispersions par le vent, au moins pour la plupart (Brown & Hovmoller 2002, Lumbsch 2008). Nos connaissances sur la structure génétique de nombreux cas de champignons ont bien sûr évolué depuis, notamment grâce à la génétique des populations, et l'emploi des reconstructions d'arbres phylogénétiques est alors justifié dans de nombreux cas (par exemple O'Donnell *et al.* 2000, Banke & McDonald 2005, Milgroom *et al.* 2008 chez des agents pathogènes émergents).

La limite commune des méthodes précédentes pour l'étude des invasions biologiques est qu'elles ne permettent pas de prendre en compte explicitement les fluctuations démographiques des populations qui ont pu avoir un impact sur le polymorphisme des marqueurs. Des méthodes plus récentes, basées sur la théorie de la coalescence, peuvent s'avérer un complément précieux lorsqu'elles sont applicables (Hey & Machado 2003, Grünwald & Goss 2011). Différents types de marqueurs, microsatellites comme séquences, peuvent servir de support à ces analyses (Brumfield *et al.* 2003, Excoffier & Heckel 2006). Ces approches reposent sur des simulations basées sur le coalescent, que l'on peut effectuer sous divers modèles d'évolution et divers modèles démographiques, et à partir de distributions *a priori* des paramètres associés. Des méthodes basées sur des MCMC (« Markov Chain Monte Carlo ») ou sur l'ABC (« Approximate Bayesian Computation ») permettent d'en tirer des distributions postérieures afin d'estimer les paramètres des modèles démographiques considérés (par ex., Beaumont 1999, Bertorelle *et al.* 2010, Csilléry *et al.* 2010, Estoup & Guillemaud 2010, Guillemaud *et al.* 2010). L'ABC, apte à traiter des modèles de plus en plus complexes, va même plus loin. Il permet de comparer différents scénarios en compétition sur la base de leurs probabilités postérieures, et estimer ainsi les erreurs statistiques associées (Estoup & Guillemaud 2010, Guillemaud *et al.* 2010). On peut ainsi tester formellement des hypothèses générées par les analyses descriptives, alors que ces dernières ne pouvaient qu'explorer l'adéquation de leurs résultats aux hypothèses formulées préalablement. Ainsi ces méthodes ont déjà été employées avec succès pour reconstruire des scénarios d'invasion (par ex. Lombaert *et al.* 2011 chez la coccinelle *Harmonia axyridis*), mais rarement chez des pathogènes émergents (mais voir Barres *et al.*, en révision, Dutech *et al.* sous presse). Ces



avancées méthodologiques deviennent applicables à de nombreux pathogènes grâce à leur extension aux espèces à biologie moins classique, notamment haploïdes (Cornuet *et al.* 2010) ou à reproduction asexuée. Elles devraient donc se multiplier chez les pathogènes fongiques et apporter d'importants compléments aux études précédentes.

### **II.2.3. Les questions supplémentaires posées par les maladies émergentes fongiques de cultures**

La biologie des champignons pathogènes de plantes n'est pas toujours bien élucidée, et l'on peut facilement se retrouver face à des obstacles spécifiques lorsqu'on envisage de reconstruire l'histoire des populations chez d'un champignon pathogène qui a émergé récemment.

Tout d'abord la notion même de population n'est pas forcément toujours évidente au premier abord, comme elle peut l'être chez des animaux ou des plantes, au vu des symptômes observables sur les plantes. La génétique des populations est alors précieuse pour préciser le cycle et les modes de reproduction et caractériser les populations. Les caractéristiques démographiques de ces organismes, restent par contre souvent méconnues. Les tailles réelles de populations peuvent paraître gigantesques au vu de la densité des symptômes observables dans un champ et des millions de spores qui peuvent être libérées. On peut s'attendre à ce que les tailles efficaces de population, *i.e.*, taille d'une population idéale de même niveau de dérive génétique, soient de même très importantes, mais ce que l'on observe au champ n'en n'est pas forcément un bon reflet. Mais il existe encore peu d'études ayant mis en œuvre des approches permettant d'estimer des paramètres démographiques. Chez *Mycosphaerella graminicola* sur blé, Zhan *et al.* (2004) estiment cependant des tailles efficaces de population de l'ordre de 24000, dépassant les ordres de grandeurs moyens de tous les autres organismes. Connaissant mal la démographie de ces organismes même en populations établies, nous n'avons *a fortiori* aucune idée de l'intensité des événements démographiques qui ont pu toucher les populations envahissantes au cours de leur histoire. De plus, dans le cas des pathogènes de grandes cultures comme les céréales, d'autres facteurs que l'histoire stochastique de dissémination sont susceptibles d'influencer les fluctuations de tailles de populations, notamment les caractéristiques biologiques des plantes hôtes, annuelles ou pérennes (Khush 2001, Barrett *et al.* 2008), mais également les pratiques agricoles (Couch *et al.* 2005, Stukenbrock *et al.* 2007, Munkacsy *et al.* 2008, Zaffarano *et al.* 2008, Gladieux *et al.* 2010).



Une autre question souvent obscure, encore liée à la biologie du pathogène, concerne la dispersion de ces organismes. Elle est généralement supposée être assurée par des spores transportées par le vent à grandes distances (Brown & Hovmoller 2002). La reconstruction de l'histoire des populations doit certes nous aider à démêler les contributions relatives des transports anthropiques et des modes de dissémination naturels du pathogène, mais encore faut-il pour cela avoir une bonne connaissance des capacités de dispersion naturelles du pathogène considéré. Des études épidémiologiques (par exemple suivis de gradients de maladies, piégeages des quantités de spores dans l'air) se sont souvent attelées à ces questions (par ex. Fouré 1984, Stover 1980, Burt et al. 1997), mais il reste complexe de déterminer précisément les distances auxquelles des propagules peuvent migrer et donner lieu à de nouveaux symptômes.

Enfin, une inconnue majeure concerne souvent le centre d'origine des espèces d'agents pathogènes envahissants. Délimiter aires d'origine et aires d'introduction est pourtant fondamental pour reconstruire les chemins d'une invasion. En l'absence de connaissance préalable sur l'espèce incriminée, plusieurs hypothèses simples sont souvent considérées mais elles ne couvrent pas tous les cas d'invasions et peuvent être sources d'erreur.

- La première hypothèse souvent avancée est que l'aire d'origine de l'espèce devrait correspondre au centre de diversité. L'hypothèse sous-jacente est que le processus d'introduction dans de nouvelles aires, généralement accompagné de goulots d'étranglements, aura réduit la diversité dans les aires envahies. En fait le centre d'origine correspondra bel et bien à un centre de diversité primaire, mais la biologie des invasions nous montre bien qu'il peut aussi se former des centres de diversité secondaires, parfois même plus diverses, suite à des cas d'introductions multiples et de mélange à partir de sources diverses (« admixture »). Par exemple Mexico serait le centre de diversité génétique de *Phytophthora infestans* mais serait en réalité un centre de diversité secondaire, tandis que l'origine serait dans les Andes (Stukenbrock & McDonald 2008). De plus dans le cas des agents pathogènes de grandes cultures comme les céréales, des expansions rapides des tailles de populations liées aux pratiques culturales peuvent aussi générer rapidement une grande diversité génétique (Stukenbrock *et al.* 2007, Zaffarano *et al.* 2008, Gladieux *et al.* 2010).

- Deuxièmement dans un cas favorable, l'aire d'origine de l'agent pathogène peut correspondre à l'aire d'origine de la plante hôte (celle-ci étant souvent mieux connue, du fait de l'importance économique des plantes considérées). On peut se baser sur l'histoire de l'hôte et de sa domestication et faire l'hypothèse simple que le pathogène aurait





coévolué avec son hôte au cours de la domestication de l'hôte, donc au sein de l'aire native de l'hôte (avant que l'hôte domestiqué ne commence à être dispersé lui-même), ce qui correspond au scénario de poursuite d'hôte (« host-tracking ») proposé par Stukenbrock et McDonald (2008). Ce mécanisme a été mis en évidence par exemple chez le pathogène *Mycosphaerella graminicola*, dont l'émergence coïncide avec la domestication du blé dans le Croissant Fertile il y a environ 10000 ans (Stukenbrock *et al.* 2007). *M. graminicola* aurait ensuite « suivi » l'introduction du blé cultivé sur les autres continents. Cependant, en raison des autres mécanismes susceptibles d'expliquer l'origine des pathogènes (Stukenbrock & McDonald 2008), la correspondance entre aire native de l'hôte et du pathogène en sera pas toujours vérifiée, par exemple dans des cas de saut d'hôte, qui peuvent se produire en dehors de la zone de domestication. C'est le cas chez le pathogène *Rhynchosporium secalis*, dont l'émergence fait suite à un saut d'hôte depuis ses hôtes sauvages sur l'orge et le seigle cultivés, en dehors de leur zone de domestication (Zaffarano *et al.* 2008). Ces phénomènes peuvent de plus être fréquents en raison, favorisé par les efforts continus d'amélioration végétale qui se sont poursuivis au-delà des premières domestications, impliquant de nombreux échanges de matériel végétal potentiellement infecté.

#### **II.2.4. Quels scénarios d'invasion caractérisent les pathogènes fongiques émergents?**

La multiplication récente des études de reconstructions de chemins d'invasion chez les champignons pathogènes envahissants permet déjà de dresser les scénarios d'introductions généralement mis à jour chez les pathogènes (Engelbrecht *et al.* 2004; Banke & McDonald 2005; Stukenbrock *et al.* 2006; Raboin *et al.* 2007; Barres *et al.* 2008; Gladieux *et al.* 2008; Høvmøller *et al.* 2008; Stukenbrock & McDonald 2008; Zaffarano *et al.* 2009; Brewer & Milgroom 2010). Les introductions multiples se révèlent ainsi omniprésentes dans les scénarios d'invasion, ce qui est vrai aussi de manière générale (Kolbe *et al.* 2004, Frankham 2005, Novak 2007, Dlugosch & Parker 2008). Selon Novak (2007) elles seraient la règle chez les plantes. Ainsi des cas d'introductions multiples ont été décelés chez un certain nombre de champignons phytopathogènes émergents à différentes échelles géographiques comme par exemple chez les pathogènes du blé *Phaeosphaeria nodorum* et *Mycosphaerella graminicola* (Banke & McDonald 2005, Stukenbrock *et al.* 2006), du pommier (*Venturia inaequalis*, Gladieux *et al.* 2008,) ou du châtaigner (Dutech *et al.* 2008, sous presse). Mais les cas d'introductions de lignées uniques ont également été rapportés comme chez le l'agent du



chancre coloré du platane *Ceratocystis fimbriata* f. *platani* à l'échelle du continent Européen (Engelbrecht *et al.* 2004), ou encore chez l'agent de la rouille de la canne à sucre *Ustilago scitaminea* à échelle mondiale (Raboin *et al.* 2007).

La reconstruction des populations sources des introductions montre qu'elles ne se situent pas toujours au sein de l'aire native, mais qu'on peut avoir des populations envahissantes servant de sources « relai ». Par exemple, la dissémination globale de la tavelure du pommier causée par *Venturia inaequalis*, qui a suivi de près celle de son hôte, proviendrait des populations européennes, elles-mêmes introduites depuis longtemps à partir de l'aire d'origine d'Asie Centrale (Gladieux *et al.* 2010). Ces exemples pourraient ainsi correspondre à des cas documentant des scénarios « tête de pont » (Lombaert *et al.* 2010), soulevant la question du rôle des évolutions adaptatives potentielles qui auraient pu avoir lieu à une seule reprise de façon localisée, avant de générer des épidémies à grande échelle.

Quels que soient les scénarios décrits, de fréquents goulots d'étranglement ont été suggérés, plus ou moins drastiques selon les espèces. De plus, suite aux introductions multiples répandues, on pourrait s'attendre à ce que les mélanges de populations soient des événements fréquents lors de ces invasions (Estoup & Guillemaud 2010). Pourtant de tels cas ont rarement été montrés chez des agents pathogènes fongiques émergents (mais voir les exemples du champignon pathogène du Pin *Ophiostoma ips*, Zhou *et al.* (2007), ou des lignées clonales françaises de *Phytophthora infestans*, Montarry *et al.* (2010)). Plus généralement dans les invasions biologiques, quelques études seulement ont proprement démontré des mélanges de populations, mais ils sont suggérés par un nombre croissant d'études (Kolbe *et al.* 2004, Lavergne & Molofsky 2007, Facon *et al.* 2008, Keller & Taylor 2010, Lombaert *et al.* 2011).

### **II.3. Etablissement et expansion des populations : déterminer les caractères clés de l'interaction entre le pathogène et l'environnement**

Dans un deuxième temps, on s'intéresse donc aux traits phénotypiques des espèces envahissantes liés à leur valeur sélective (i.e., nombre de descendants participant effectivement à la génération suivante) dans l'environnement. Cela requiert premièrement de cibler et décrire les traits sur lesquels se base l'adéquation entre populations envahissantes et environnements envahis. Certains traits phénotypiques pourront être spécifiques aux systèmes étudiés, comme par exemple une certaine tolérance climatique. Mais plus généralement, il sera crucial de s'intéresser aux traits d'histoire de vie qui influencent fortement la valeur



sélective chez tous les organismes, puisqu'ils régulent directement le rythme de reproduction et la quantité de descendants (Stearns 1992).

### **II.3.1. Importance des traits d'histoire de vie liés à la colonisation et à la compétition lors des invasions**

Une pensée historiquement omniprésente est l'hypothèse que les espèces envahissantes devraient être de bons colonisateurs (Lodge 1993, Kolar & Lodge 2001, Duyck *et al.* 2007, au sens de la stratégie r, définie historiquement par une série de traits qui leur confèrent un fort taux de croissance instantanée de population (Reznick 2002): temps de génération court, phase juvénile courte, effort reproducteur élevé avec production de descendance petite, nombreuse et très mobile (conférant de fortes capacités de dispersion). Cela a d'ailleurs été rapporté notamment des plantes (Lachmuth *et al.* 2011) comme l'exemple bien documenté de conifères envahissants (Richardson & Rejmanek 2004). Ils seraient avantagés dans des environnements non compétitifs tels que les premiers stades de succession végétale ou des environnements instables, au contraire de la stratégie plus compétitive K qui suppose une reproduction tardive et un faible nombre de descendants de grande taille. Cette théorie est basée sur l'hypothèse d'un compromis évolutif (« trade-off ») entre traits qui empêche de maximiser le taux de croissance de population en même temps que la compétitivité (Reznick 2002, Kneitel & Chase 2004), dans des contextes où les ressources sont limitées et où l'on ne peut donc pas maximiser leur allocation à tous les aspects de l'histoire de vie. En réalité cette vision a été complexifiée et enrichie par la théorie de l'évolution de l'histoire de vie (Reznick 2002) qui a revisité les stratégies r et K (avec notamment l'hypothèse du compromis compétition – colonisation). Mais ces notions restent utilisées dans la littérature récente et opérationnelles dans le cadre de notre discussion (Duyck *et al.* 2007). Quoi qu'il en soit, la plupart de ces prédictions concernent l'arrivée dans des environnements vierges de tout compétiteur.

Or, bien qu'encore peu étudiée dans la littérature des invasions biologiques, la compétition est reconnue comme une composante importante de l'environnement biotique d'une espèce envahissante (Byers 2000, Evans 2004, Vila & Weiner 2004) et l'on peut s'attendre alors à ce que les prédictions précédentes soient modifiées car la théorie des traits d'histoire de vie prévoirait alors un avantage à la stratégie compétitive K (Duyck *et al.* 2007, Burton *et al.* 2010). C'est en effet ce que documentent des études sur des invasions successives d'escargots d'eau douce en Martinique (Facon *et al.* 2006, 2008), ou de mouches



des fruits à La Réunion (Duyck *et al.* 2007). La compétition est particulièrement importante dans le cas des émergences de pathogènes de cultures, car il n'est pas rare que certains pathogènes émergents, souvent responsables de sévères épidémies, envahissent de nouveaux territoires au détriment de souches ou espèces proches qui s'étaient établies préalablement. Cela a en fait peu été documenté au niveau interspécifique, mais de nombreux cas de remplacement s'observent entre souches ou pathotypes d'une même espèce, comme cela a fréquemment été le cas chez *Phytophthora infestans* que ce soit en Colombie (Miller *et al.* 1998), en Europe (Day & Shattock 1997). Cette problématique constitue même l'une des questions centrale en épidémiologie qui cherche à comprendre quels facteurs influencent la persistance et l'invasion des pathogènes au sein d'un agrosystème, et notamment pourquoi certains agents pathogènes déclinent, d'autres coexistent et d'autres encore supplantent les précédents (Gilligan & van den Bosch 2008). Pourtant, même chez les pathogènes, comprendre les traits d'histoire de vie clé qui confèrent aux nouveaux pathogènes leur avantage compétitif et donc leurs succès d'invasion, reste une question encore peu étudiée sous cet angle.

### **II.3.2. Réinvestir les méthodes théoriques et expérimentales de l'épidémiologie**

L'épidémiologie des maladies de plantes est une discipline qui se penche sur la compréhension des facteurs affectant les dynamiques de maladies dans l'espace et dans le temps (Milgroom & Peever 2003). Ainsi elle tient compte simultanément des populations de pathogènes et de plantes hôtes dans un contexte environnemental. De nature interdisciplinaire, elle doit donc faire appel à l'alliance des modèles théoriques et de l'expérimentation.

Une approche expérimentale est donc premièrement nécessaire afin de mesurer, en conditions contrôlées, des traits d'histoire de vie liés à la valeur sélective des pathogènes dans l'environnement considéré. L'environnement envahi, dans le cas des pathogènes est représenté par les cultures hôtes, donc « vivant » même s'il n'évolue pas ou très lentement au sein des agro-systèmes industriels. Ainsi les traits d'histoire de vie sont aussi des traits d'infection, caractéristiques de l'interaction hôte-pathogène-environnement (Barrett *et al.* 2008). Chez les champignons pathogènes, les traits d'histoire de vie principaux susceptibles d'influencer la valeur sélective seront notamment la quantité de spores produite ou encore la période de latence (durée entre le début de l'infection et le début de la sporulation), (Pariaud *et al.* 2009). De telles mesures ont en fait rarement été abordées sous cet angle chez ces





organismes. Pourtant une importante littérature en phytopathologie s'intéresse depuis longtemps à mesurer ces traits, mais sous l'angle de composantes d'agressivité plutôt que de traits d'histoire de vie (par ex. Miller *et al.* 1998, Carlisle *et al.* 2002, Pariaud *et al.* 2009b, Caffier *et al.* 2010). L'agressivité est à la base une notion opérationnelle en phytopathologie, décrivant la quantité de dégâts et la vitesse à laquelle ils se propagent lors d'une épidémie. Elle s'intéresse donc à l'interaction entre l'hôte et le pathogène et permet par exemple de mener des études de comparaison de résistance entre différentes variétés hôtes. Si l'agressivité peut être mesurée directement par la sévérité globale des symptômes, elle est également souvent décomposée en différentes composantes d'agressivité que sont la période d'incubation, la latence ou encore la quantité de spores produites (Pariaud *et al.* 2009a). C'est là que se rejoignent les notions d'agressivité et de valeur sélective, bien qu'elles ne se recoupent pas totalement. Un certain nombre de composantes d'agressivité s'avèrent correspondre aux traits reliés à la valeur sélective qui nous intéressent ici (Pariaud *et al.* 2009a). Mais là encore, les cycles complexes de la plupart des champignons phytopathogènes peuvent rapidement compliquer non seulement les mesures en elles-mêmes d'un point de vue technique, mais aussi notre compréhension des traits qui peuvent s'avérer les plus importants dans le succès d'établissement des pathogènes, par exemple leur supériorité compétitive. Chez des pathogènes alternant divers modes de reproduction, il ne sera pas forcément simple de savoir au premier abord quels traits liés à tel ou tel mode de reproduction ont le rôle le plus crucial pour la valeur sélective du pathogène envahissant, par exemple sa supériorité compétitive sur un autre pathogène résident.

Ainsi, explorer comment différents traits d'histoire de vie se combinent pour influencer la valeur sélective, requiert l'apport des modèles théoriques. On peut alors réinvestir les concepts d'invasion et persistance des pathogènes explorés par les modèles épidémiologiques. Ces modèles décrivent généralement la dynamique spatio-temporelle des différents compartiments sains et infectés de l'hôte (Gilligan & van den Bosch 2008), où les probabilités de transitions d'un compartiment à l'autre sont directement liées aux traits d'histoire de vie du pathogène décrivant sa croissance, sa reproduction, sa dispersion. Les traits d'histoire de vie mesurés expérimentalement peuvent alors permettre de paramétrer les modèles (Montarry *et al.* 2010). Des mesures de valeur sélective pourront alors être dérivées de ces modèles en fonction de ces paramètres biologiques (Gilchrist *et al.* 2006). La compétition peut notamment être intégrée à de tels modèles, afin d'étudier comment les traits des différents compétiteurs peuvent jouer sur leur valeur sélective relative et donc sur l'issue de la



compétition. Cependant ces questions restent très peu explorées chez les agents pathogènes de plantes et nécessiteront d'adapter les modèles existants au cas spécifique d'organismes à cycles complexes, alternant parfois divers modes de reproduction et différentes plantes hôtes.

#### **II.4. Vers l'intégration de l'histoire démographique des populations et de leurs stratégies de vie pour décrypter le succès d'invasion**

Dès lors que l'on a pu cibler quels sont les traits d'histoire de vie clé pour l'établissement d'une part, et que l'on a retracé l'histoire des populations envahissantes depuis leurs populations sources d'autre part, on est alors en mesure d'adresser plus précisément des questions clé dans le domaine des invasions biologiques, au centre de nombreuses recherches actuelles : quel est le rôle de la diversité génétique dans le succès d'invasion ? (Dlugosch & Hays 2008, Dlugosch & Parker 2008, Olivieri 2009, Crawford & Whitney 2010, Eales *et al.* 2010). En effet, outre leurs effets démographiques immédiats, goulots d'étranglement et mélanges de populations conditionnent directement la diversité génétique des populations, et donc peuvent avoir un impact important que leur potentiel évolutif, et notamment adaptatif. Notamment quel impact peuvent avoir les goulots d'étranglement sur le potentiel adaptatif des populations envahissantes et leur succès d'établissement ? Par ailleurs, est-ce que le mélange de populations, favorisant la création des nouvelles combinaisons génotypiques pourrait favoriser le succès d'invasion des populations ?

On sait que la diversité moléculaire observée aux marqueurs neutres n'est pas forcément une bonne représentation de la diversité au niveau des traits quantitatifs. Ainsi, des mesures expérimentales de ces traits devront faire l'objet d'approches comparatives entre populations, afin d'étudier s'ils ont évolué de manière adaptative pour permettre l'invasion. L'étude de ces questions passe par exemple par des approches de génétique quantitative comparant traits d'histoire de vie des populations envahissantes à ceux des populations sources, ou encore des approches dites  $Q_{st}/F_{st}$  comparant les degrés de divergence aux loci neutres et aux loci affectant des traits phénotypiques, afin de déceler des divergences générées spécifiquement par les pressions de sélection (Keller & Taylor 2008). Jusque là, quelques études ont pu montrer par exemple que les goulots d'étranglement peuvent en effet affecter négativement le potentiel adaptatif (Pujol & Pannell 2008), d'autres études ayant cependant montré qu'ils ne prévenaient pas forcément des adaptations évolutives rapides (Dlugosch & Parker 2008), voire pourraient même être bénéfique au succès d'invasion en purgeant les



allèles délétères qui affectent la valeur sélective des populations de l'aire d'origine (Facon *et al.* 2011). Des effets positifs du mélange de populations sur la performance des populations envahissantes dans l'environnement ont également pu être démontrés dans quelques études (Lavergne & Molofsky 2007, Wolfe *et al.* 2007, Facon *et al.* 2008, 2011). A l'heure actuelle ces questions demandent encore à être mieux documentées, et notamment chez les pathogènes émergents chez qui elles ont très peu été explorées. Il reste notamment beaucoup d'hypothèses à tester dans le domaine des mécanismes évolutifs liés aux invasions, telles que celles liées à l'évolution de la dispersion par exemple, ou à l'évolution d'une balance entre différents modes de reproduction sexués et asexués, qui pourraient être étudiées de façon spécifique chez ces organismes.

### **III. La maladie des raies noire du bananier causée par *Mycosphaerella fijiensis* : un modèle de maladie émergente fongique à l'échelle mondiale**

#### **III. 1. Impacts en agriculture**

Au quatrième rang économique mondial des cultures destinées à l'alimentation après le riz, le blé et le maïs (Arias *et al.* 2003), la banane est cultivée dans plus de 130 pays à travers le monde, principalement au sein des régions tropicales et subtropicales du Sud-Est asiatique, du Pacifique, d'Afrique et d'Amérique. La production mondiale est de 105 millions de tonnes (FAO, 2006) dont 16 millions sont destinées à l'exportation. Les cultivars modernes, parthénocarpiques et stériles, sont issus majoritairement de croisements intra et interspécifiques entre les espèces diploïdes sauvages *Musa acuminata* et *Musa balbisiana* (Simmonds 1966). La domestication des bananiers a débuté il y a plusieurs milliers d'années dans le Sud-Est Asiatique et le Pacifique, aire d'origine des bananiers où l'on trouve une grande variété d'espèces sauvages (Simmonds 1962, Perrier *et al.* 2011). Leur dissémination sur les autres continents serait très ancienne également (Perrier *et al.* 2011). On distingue deux grands types de bananes d'un point de vue alimentaire : les bananes douces, dites « dessert », et les bananes à cuire comprenant notamment les bananes plantains. Ressource alimentaire majeure dans la majorité des régions productrices, les bananes sont ainsi cultivées au sein des jardins et des plantations paysannes destinées à l'autoconsommation et au commerce local, ce qui représente environ 85% de la quantité totale mondiale produite (Arias



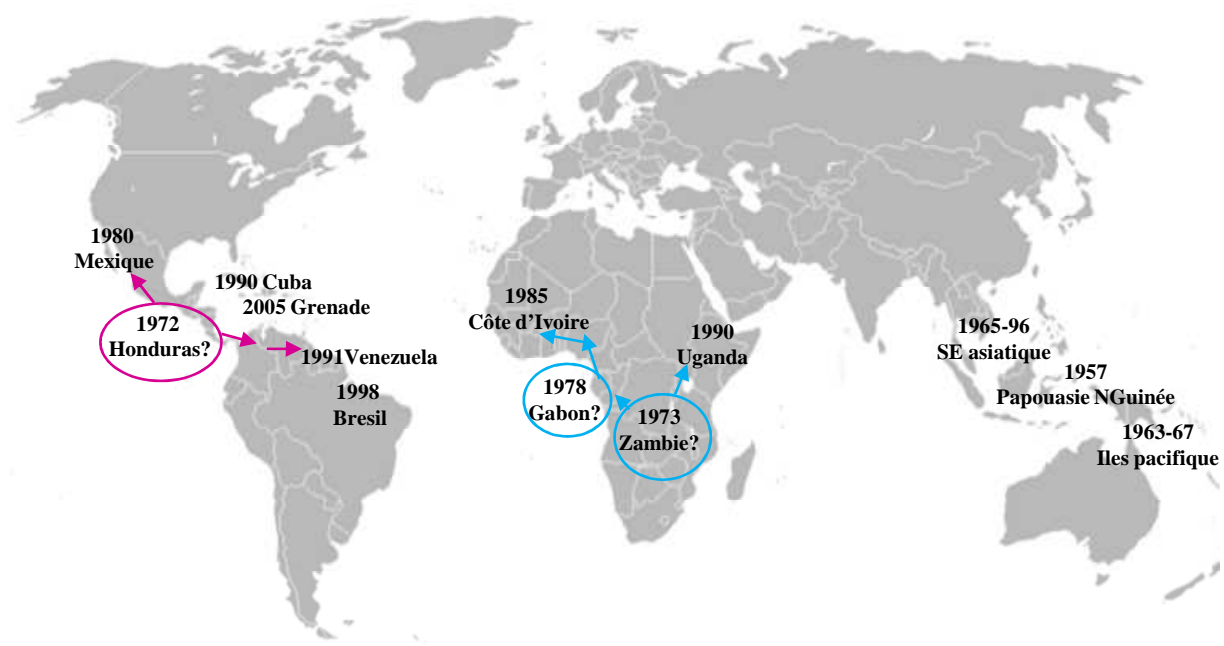
**Figure 2: Sévère défoliation causée par une cercosporiose du bananier.** (Photos : collection de L. Lapeyre de Bellaire)

*et al.* 2003). D'autre part, un certain nombre de pays producteurs sont aussi d'importants exportateurs, cultivant ainsi la banane en monoculture au sein de vastes plantations industrielles. Cette production vouée à l'exportation, en grande partie vers l'Europe et les Etats-Unis, concerne surtout la banane dessert, mais une part croissante des plantains (Marin *et al.* 2003, Churchill *et al.* 2010).

Parmi les divers parasites auxquels doivent faire face ces cultures, la maladie des raies noires, causée par le champignon pathogène *Mycosphaerella fijiensis*, est la plus destructrice maladie du feuillage (Figure 2), (Stover & Simmonds 1987). Elle affecte notamment tous les cultivars communs, bananes et plantains. Décrite pour la première fois à Fidji en 1963 (Rhodes 1964), la maladie des raies noires du bananier (BLSD pour Black Leaf Streak Disease ou Black Sigatoka) a émergé sur tous les continents en l'espace de deux décennies, et apparaît comme la maladie des bananiers et plantains la plus coûteuse d'un point de vue économique (Churchill *et al.* 2010). Le champignon *M. fijiensis* affecte de façon directe les feuilles des bananiers où il provoque lésions et nécroses susceptibles de détruire complètement le feuillage des cultivars les plus sensibles en l'absence de mesures de contrôle. Outre de sérieuses pertes de rendement dues à la réduction de la surface photosynthétique, d'autres pertes concernent la production vouée à l'exportation. La maladie provoque une maturation prématurée des bananes produites, impact très dommageable par les filières d'exportation pour lesquelles les fruits récoltés verts doivent rester vert (durée de vie verte) pendant au moins 2 à 3 semaines, durée nécessaire au transport de fruits des zones de productions vers les zones d'importation (Marin *et al.* 2003, Abadie *et al.* 2008, de Lapeyre de Bellaire *et al.* 2010). La sélection de variétés partiellement résistantes de façon durable, est donc un important sujet de recherches. A l'heure actuelle la production industrielle de bananes desserts, principalement des cultivars sensibles du sous-groupe Cavendish, requiert l'utilisation intensive de fongicides chimiques (Marin *et al.* 2003, de Lapeyre de Bellaire *et al.* 2010). Outre leurs effets néfastes bien connus sur l'environnement, ceux-ci sont particulièrement nocifs pour les cultivateurs et l'ensemble des populations vivant juste à proximité des agro-systèmes industriels.

### **III. 2. Suivi historique de l'émergence mondiale**

La première description de la maladie à Fidji en 1963 (Rhodes 1964) a fait suite à une épidémie particulièrement sévère apparue dans une grande plantation commerciale particulièrement bien surveillée. Les prospections qui ont suivi indiquent que la maladie



**Figure 3: Carte très synthétique des données de premières détection de la maladie dans quelques pays, avec lieux d'introductions et chemins de d'invasion suspectés historiquement (repérés par un cercle).**

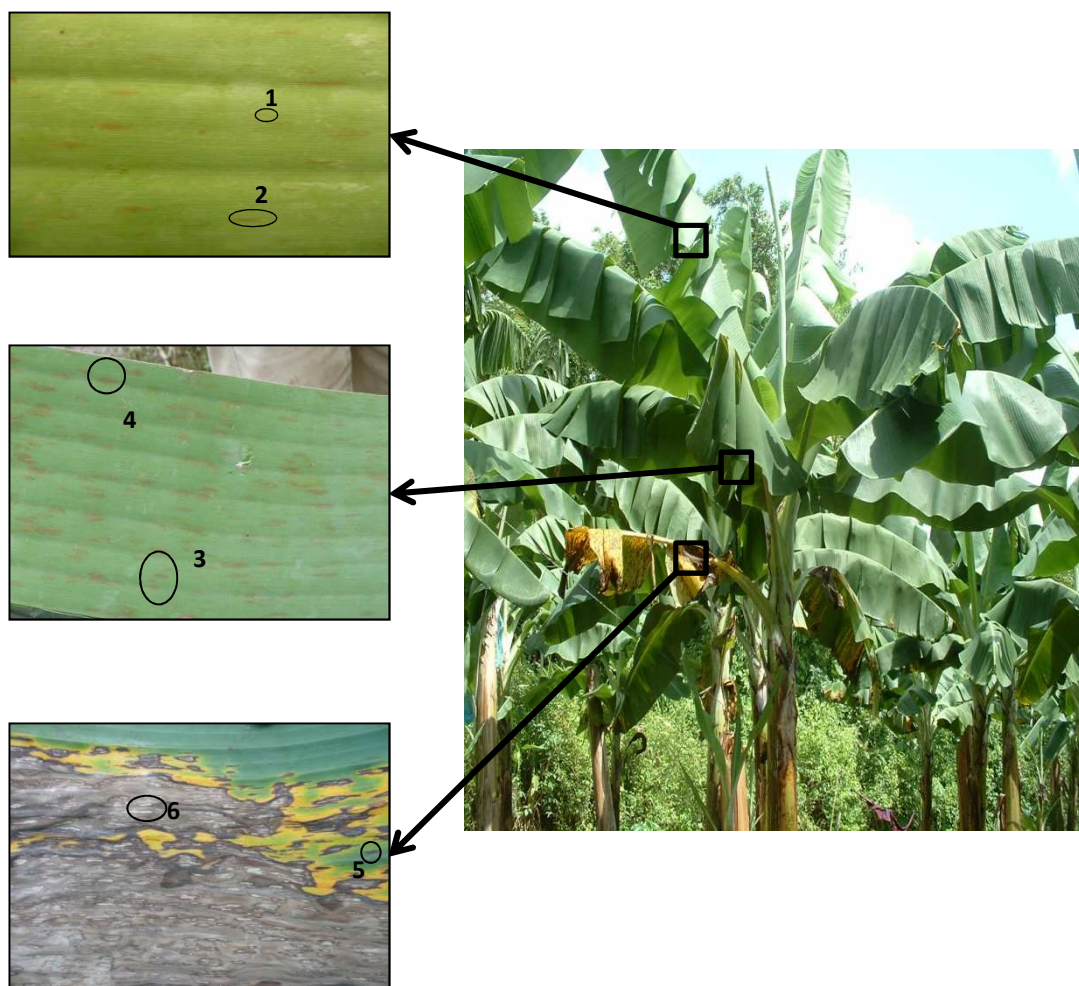


devait être présente dans certaines régions du Sud-Est Asiatique bien avant cette première détection, depuis au moins les années 50 voire les années 30, mais sous une forme peut-être moins sévère donc moins facilement détectable (Stover 1978). Ses symptômes sont de plus très proches à l'œil nu de ceux de la maladie de Sigatoka (ou Sigatoka jaune), causée par l'espèce apparentée *M. musicola*, largement présente au sein du Sud-Est Asiatique. Les premières détections de la maladie des raies noires semblent ainsi avoir coïncidé avec les modifications modernes de pratiques agricoles, notamment l'implantation des premières grandes plantations industrielles dans l'aire Pacifique (Rhodes 1964).

Suite à cela, *M. fijiensis* a émergé et s'est propagé en Amérique et en Afrique à partir des années 70, déclenchant ainsi de sérieuses épidémies sur tous les continents (Pasberg-Gauhl *et al.* 2000). Un certain nombre de relevés historiques relatent en partie la progression spatio-temporelle de la maladie (Figure 3), à partir d'une aire d'origine suspectée dans le Sud Est Asiatique (Stover 1978). Ces données manquent cependant de précision et de fiabilité quant aux dates réelles d'arrivée de la maladie dans une région, généralement bien antérieures aux premières détections (Jones 2000). C'est particulièrement vrai en Afrique où l'histoire d'introduction reste obscure, alors qu'elle semble avoir été mieux renseignée en Amérique (Mourichon & Fullerton 1990, Pasberg-Gauhl *et al.* 2000). A l'heure actuelle, la maladie progresse encore au sein de l'archipel des Caraïbes, où elle est présente depuis les années 90.

Cette émergence mondiale a fait suite à celle de la maladie de Sigatoka, causée par *M. musicola*, détectée dans le SE Asiatique dès le début du siècle (Zimmerman 1902, Meredith 1970), et qui s'est également répandue mondialement dans les années 30 (Mourichon 2003). Lors de son expansion dans les îles du Pacifique, en Afrique et Amérique, *M. fijiensis* a alors supplanté *M. musicola* en l'espace de quelques années (Jones 2000). Dans ces régions *M. musicola* peut cependant subsister dans quelques productions d'altitudes sur tous les continents (Mouliom Pefoura 1996, Jacome 1992). De plus *M. musicola* resterait largement présent dans le Sud Est Asiatique (Jones 2000). Les deux espèces pourraient y cohabiter dans les mêmes régions, du fait de la grande diversité de cultivars et bananiers sauvages, susceptibles de présenter des degrés de résistances variés. Toutefois, la répartition relative des deux espèces au sein du Sud Est Asiatique, et la distribution des bananiers sauvages et cultivés restent peu renseignées à notre connaissance.

### **III. 3. Biologie du pathosystème bananier/*Mycosphaerella fijiensis***

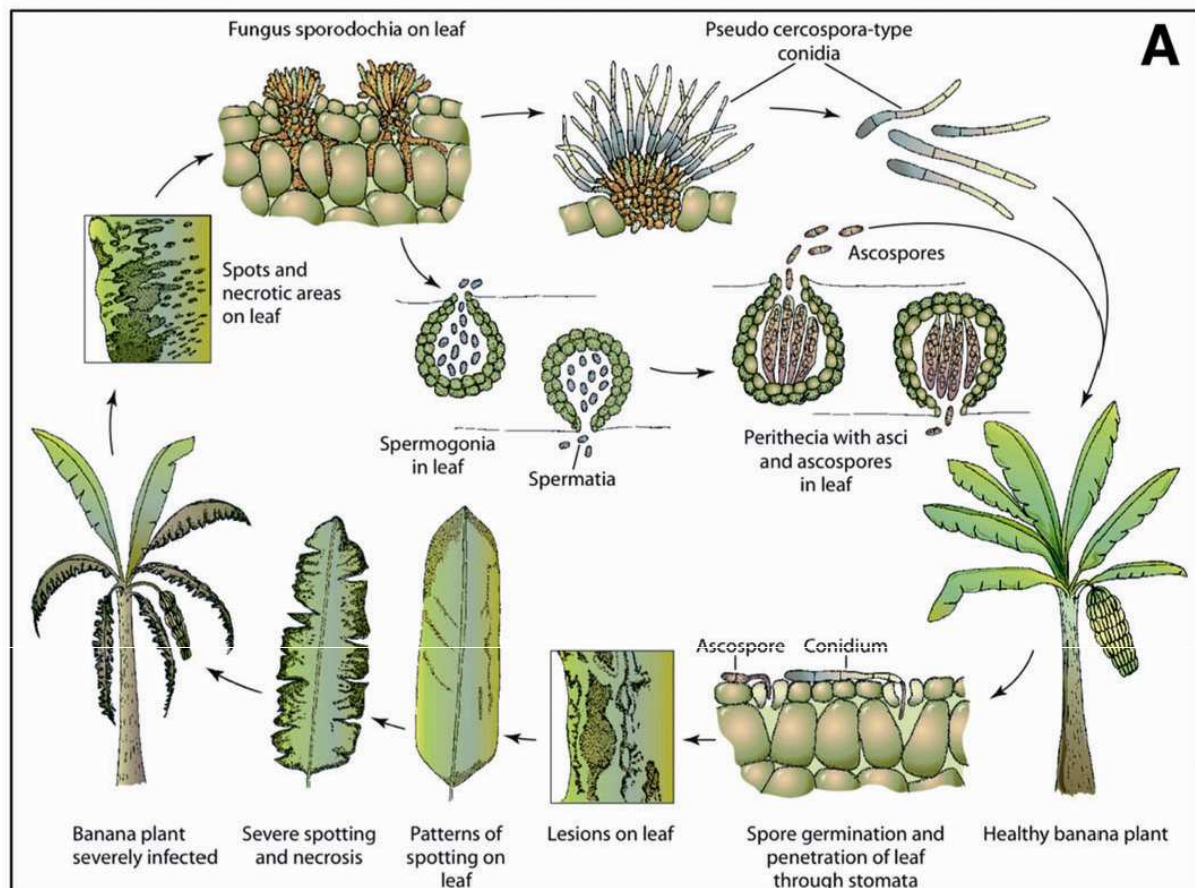


**Figure 2: Photographies illustrant les six stades d'avancement des symptômes de la maladie des raies noires (Meredith & Lawrence 1969, Fouré 1985) au niveau d'un bananier infecté. (Photos : collection de L. Lapeyre de Bellaire)**

Le bananier est une plante pérenne, dont la culture se déroule de manière continue toute l'année et qui est réceptive de façon constante à la maladie. C'est une plante monocotylédone qui émet une nouvelle feuille à son sommet de façon quasi hebdomadaire, ce qui se traduit par un décalage régulier des rangs de chaque feuille vers la base. Les jeunes feuilles nouvellement émises et encore enroulées, dites au stade « cigare », sont les plus réceptives aux spores du champignon (Gauhl *et al.* 2000). Six stades d'avancement et maturation des lésions de la maladie ont été décrits (Meredith & Lawrence 1969, Fouré 1985), depuis l'apparition des tirets sur les plus jeunes feuilles jusqu'à la coalescence des lésions puis l'expansion de plages nécrotiques sur des feuilles plus basales (Figure 4).

L'évolution des symptômes résulte donc du déroulement du cycle de vie hémibiotrophe du champignon parasite au sein du tissu foliaire. Le cycle de vie de *M. fijiensis*, champignon ascomycète haploïde, a été précédemment décrit (Figure 5). Après germination d'une spore sur une feuille et pénétration du mycélium au niveau des stomates, la croissance végétative se déroule au détriment des tissus foliaires, initiant les premières lésions. Ceci marque la fin de la période d'incubation qui correspond à la période temporelle entre l'infection et l'apparition des symptômes. Puis le champignon développe ses structures de reproduction à la surface ou à l'intérieur des feuilles. Cette fois on parle de période de latence pour désigner la période temporelle entre l'infection et le début de la reproduction (sporulation). *M. fijiensis* est un champignon hétérothallique (i.e., nécessitant la rencontre de deux thalles de types sexuels opposés pour effectuer la reproduction sexuée) qui présente les deux modes de reproduction sexué et asexué au cours de son cycle, et cela dans toute son aire de répartition. La formation des structures de reproduction asexuées (conidiophores) conduit à la production de spores (conidies) majoritairement à la surface inférieure de la feuille, au niveau des lésions de stades 2 et 3 (Fouré 1987). Plusieurs cycles asexués peuvent ainsi s'effectuer, avant que le champignon n'entame la formation des structures de reproduction sexuées (périthèces) lorsque les lésions matures de types sexuels opposés commencent à coalescer. Au sein des périthèces sont produites les ascospores, de façon beaucoup plus abondante que la production des conidies, ce qui ferait de la reproduction sexuée le mode de reproduction dominant chez *M. fijiensis* (Marin *et al.* 2003). On appelle période infectieuse la durée pendant laquelle le champignon sporule.

De manière importante, les deux types de spores sexuées et asexuées présentent des capacités de dispersion différentes. Les conidies sont principalement détachées et dispersées par le ruissellement des eaux de pluie, bien qu'une partie puisse également être détachée et



**Figure 5: Cycle de vie simplifié des champignons responsables des cercosporioses du bananier.**  
(Source : Churchill *et al.* 2010)

transportée par le vent (Fouré 1987, Burt *et al.* 1997). Elles assurent une dispersion sur de courtes distances de quelques mètres maximum, ré-infestant les feuilles adjacentes du même plant et des plants voisins (Burt *et al.* 1997, Rieux 2011). Les ascospores, d'une plus petite taille que les conidies, sont activement projetées dans l'air suite à l'explosion des périthèces et transportées par le vent. Il a été montré que les ascospores sont ainsi dispersées sur plusieurs centaines de mètres voire quelques kilomètres (Rieux 2011, Abadie & Bonnot *in prep.*). Il a de plus été suggéré qu'elles pourraient assurer des événements rares de dispersion à beaucoup plus grande distance (Burt 2003) mais sans démonstration claire de ce phénomène. Leur sensibilité aux rayons UV (Parnell *et al.* 1998) préviendrait toute possibilité d'infestation sur des distances de plusieurs centaines de kilomètres. Les ascospores sont ainsi considérées comme la source principale d'inoculum permettant le maintien et la progression spatiale de la maladie (Stover 1980, Burt 2003, Marin *et al.* 2003).

Enfin, un autre moyen majeur de dispersion de la maladie, serait le transport de matériel végétal infecté via les activités humaines et échanges commerciaux. Ceux-ci auraient été largement répandus historiquement, que ce soit sous la forme de feuilles de bananiers infectées servant à envelopper des marchandises ou d'échanges de rejets destinés à être replantés (Stover 1978, Perrier *et al.* 2010).

### **III. 4. Structure et diversité génétique des populations**

Outre ces données historiques et phytopathologiques, quelques études de génétique des populations ont été conduites précédemment, apportant des informations sur la structure et la diversité des populations chez ce pathogène. Nous considérons qu'au champ, chaque individu correspond à une lésion isolée. Il est à noter que ces études ont été menées sur des individus - ou isolats - issus d'ascospores, et ne décrivent donc que la fraction des populations issue de reproduction sexuée. Les premières études, basées sur des marqueurs RFLP et/ou des marqueurs microsatellites, ont ainsi pu montrer que les populations sont caractérisées par le maintien de forts niveaux de diversité génétique et une panmixie (brassage aléatoire des gamètes) sur une échelle allant de la plante-hôte à la plantation au moins (Hayden *et al.* 2003, Rivas *et al.* 2004).

A des échelles supérieures (entre localités, pays et continents), de forts niveaux de différenciation inter-populationnelle ont également été montrés, et pourraient résulter d'événements de fondation lors de l'expansion globale de la maladie (Rivas *et al.* 2004). De plus, Halkett *et al.* (2010) ont montré des patrons d'isolement par la distance le long de



transects de plusieurs centaines de kilomètres au niveau de bassins de production au Cameroun et au Costa Rica. Ces résultats suggèrent ainsi l'existence de fronts d'expansion graduels de la maladie au sein de continents. Se superposant à ces patrons réguliers, des signatures d'évènements plus stochastiques ont également été décelées, comme l'existence d'une discontinuité génétique abrupte ne trouvant pas d'explication géographique au Cameroun (Rieux *et al.* 2011).

Enfin, de plus forts niveaux de diversité génétique ont de plus été montrés dans quelques zones échantillonnées du Sud-Est Asiatique, notamment en Papouasie Nouvelle Guinée et Philippines (Carlier *et al.* 1996, Hayden *et al.* 2003), renforçant l'hypothèse d'une origine Sud-Est Asiatique du pathogène, qui aurait co-évolué avec sa plante hôte dans son aire d'origine (Stover 1978., Carlier *et al.* 1996).

Finalement, parce-que l'ensemble de ces études génétiques se sont basées sur des marqueurs différents, développés pour travailler à des échelles géographiques variées, elles ne peuvent être directement comparées ni intégrées en un panorama global de l'histoire d'émergence de *M. fijiensis*.

#### **IV. Objectifs de la thèse**

La maladie des raies noires nous apparaît comme un cas caractéristique de maladie émergente mondiale de grandes cultures à forts enjeux d'études fondamentales et appliquées. Compte tenu de la dissémination très récente de *M. fijiensis* comparée à l'implantation globale de sa plante hôte, nous nous trouvons précisément dans un cas d'invasion biologique fongique en agriculture. Cette thèse propose donc de progresser dans la compréhension des fondements écologiques et évolutifs de l'invasion mondiale de *M. fijiensis*, permettant aussi de documenter les mécanismes d'invasion des pathogènes fongiques de grandes cultures d'une manière plus générale. Nous nous sommes attachés à répondre à deux grands objectifs qui s'inscrivent dans la démarche proposée en introduction :

- ❖ Reconstruire les scénarios d'invasion de *M. fijiensis* à échelle mondiale et continentale, comprenant les itinéraires d'invasion ainsi que les évènements démographiques ayant pu affecter les populations lors de la dissémination.





Premièrement nous avons cherché à dresser un panorama mondial des routes et scénarios d'invasion de *M. fijiensis* en se basant sur l'analyse des patrons de diversité et structure génétique des populations (**Chapitre 1**). Nous avons ensuite mené des études à des échelles continentales destinées à tester plus formellement les hypothèses issues du chapitre I par des méthodes statistiques plus récentes, ainsi qu'à mieux comprendre les processus d'invasion au sein des continents (**Chapitre 2**).

- ❖ Décrire et comprendre ce qui fait l'adéquation efficace entre *M. fijiensis* et les environnements envahis.

Pour cela nous nous sommes attelés à décrire la stratégie d'histoire de vie de *M. fijiensis*, en comparaison à celle de *M. musicola*, et exploré comment les combinaisons de traits d'histoire de vie mises en évidence peuvent conférer son avantage à *M. fijiensis* (**Chapitre 3**).



# CHAPITRE 1



# CHAPITRE 1

## **Reconstitution de la propagation mondiale de *M. fijiensis***

L'objet de ce chapitre est de dresser un panorama mondial de la structure et diversité génétique des populations de *M. fijiensis*, afin de préciser l'aire d'origine du pathogène et son histoire d'invasion à l'échelle mondiale. Il apporte des réponses plus ou moins approfondies aux questions suivantes : peut-on confirmer une aire d'origine Sud-Est Asiatique ? Quelles sont les routes de dissémination globale et la localisation des populations sources ? Combien d'introductions ont initié l'invasion sur les différents continents ? Peut-on mettre en évidence les évènements démographiques qui ont accompagné ces introductions ?

Nous nous sommes basés pour cela :

- sur une collection mondiale d'isolats conservés au laboratoire BGPI issus de diverses collectes destinées à des études précédentes. Certains échantillons correspondent à des populations échantillonnées dans une même localité, mais un certain nombre d'entre eux correspondent aussi à des individus seuls ou en petit groupes dispersés à travers tous les continents ;

- sur l'emploi de jeux de marqueurs neutres communs sur tous les continents, avec le développement de nouveaux microsatellites dont la résolution permet de travailler à l'échelle mondiale. Quelques marqueurs basés sur séquences ont également été mis au point, permettant entre autre d'exploiter au maximum la collection et balayer les aires géographiques de manière la plus exhaustive possible en analysant tous les individus isolés de la collection ;

- sur une synthèse bibliographique des données historiques, afin de pouvoir confronter les hypothèses émises historiquement à celles issues de nos analyses de génétique des populations.

Ce chapitre repose sur deux articles publiés : un premier article qui constitue le cœur de ce chapitre, explorant les questions posées ci-dessus, un deuxième article technique correspondant au développement de la plupart des marqueurs microsatellites utilisés dans l'étude principale.



# Contrasting introduction scenarios among continents in the worldwide invasion of the banana fungal pathogen *Mycosphaerella fijiensis*

S. ROBERT, V. RAVIGNE, M-F. ZAPATER, C. ABADIE and J. CARLIER

CIRAD, UMR BGPI, Campus International de Baillarguet, TA A-54K, F-34398 Montpellier Cedex 5, France

## Abstract

Reconstructing and characterizing introduction routes is a key step towards understanding the ecological and evolutionary factors underlying successful invasions and disease emergence. Here, we aimed to decipher scenarios of introduction and stochastic demographic events associated with the global spread of an emerging disease of bananas caused by the destructive fungal pathogen *Mycosphaerella fijiensis*. We analysed the worldwide population structure of this fungus using 21 microsatellites and 8 sequence-based markers on 735 individuals from 37 countries. Our analyses designated South-East Asia as the source of the global invasion and supported the location of the centre of origin of *M. fijiensis* within this area. We confirmed the occurrence of bottlenecks upon introduction into other continents followed by widespread founder events within continents. Furthermore, this study suggested contrasting introduction scenarios of the pathogen between the African and American continents. While potential signatures of admixture resulting from multiple introductions were detected in America, all the African samples examined seem to descend from a single successful founder event. In combination with historical information, our study reveals an original and unprecedented global scenario of invasion for this recently emerging disease caused by a wind-dispersed pathogen.

**Keywords:** Emerging disease, microsatellite markers, networks, population genetics, sequence-based markers

Received 4 March 2011; revision accepted 23 November 2011

## Introduction

The reconstruction of historical sources and pathways of biological invasions is essential from both fundamental and applied perspectives (Estoup & Guillemaud 2010). The success of establishment and spread of invading populations is thought to result from complex demographic scenarios that often include bottlenecks and/or admixture (Estoup & Guillemaud 2010). Indeed, multiple introductions now appear as a frequent component of many invasions (Kolbe *et al.* 2004; Frankham 2005; Novak 2007; Dlugosch & Parker 2008). Thus, the detection and characterization of demographic events

associated with introductions represent an integral part of the reconstruction of invasion pathways; such tasks generally require the combination of historical data documenting early detection and progression in invaded countries with a variety of population genetic tools (Estoup & Guillemaud 2010).

Emerging plant diseases are often characterized by large geographical range expansions and can therefore be considered as bioinvasions (Anderson *et al.* 2004; Desprez-Loustau *et al.* 2007; Stukenbrock & McDonald 2008). In comparison with other organisms, only a few studies have attempted to decipher historical source populations and introduction events in invasive fungal plant pathogens at large scale (Engelbrecht *et al.* 2004; Banke & McDonald 2005; Stukenbrock *et al.* 2006; Raboin *et al.* 2007; Barres *et al.* 2008; Gladieux *et al.* 2008;

Correspondence: Jean Carlier, Fax: 33 0 4 99 62 48 48;  
E-mail: jean.carlier@cirad.fr

Hovmøller *et al.* 2008; Stukenbrock & McDonald 2008; Zaffarano *et al.* 2009; Brewer & Milgroom 2010). Contrasting results have been yielded by these studies, which generally encounter specific difficulties.

Firstly, locating invasion source populations is often complicated because species centres of origin are unknown for most fungal crop pathogens (Desprez-Loustau *et al.* 2007). Disease emergence may actually result from different evolutionary mechanisms, such as host jump, or host-tracking defined by the domestication of the pathogen along with its host plant (reviewed in Stukenbrock & McDonald 2008). In the case of host-tracking, the centres of origin of both host and pathogen should coincide. This was described for instance in *Mycosphaerella graminicola* on wheat (Stukenbrock *et al.* 2007), *Magnaporthe oryzae* on rice (Couch *et al.* 2005), *Phytophthora infestans* on potato (Gomez-Alpizar *et al.* 2007), *Ustilago maydis* on maize (Munkacsy *et al.* 2008) and *Venturia inaequalis* in apple (Gladieux *et al.* 2008, 2010), originating from areas as diverse as the Fertile Crescent, Asia, Andes, Mexico and central Asia, respectively. Population genetics analyses are always required to verify such hypotheses. They are based on the simple idea that the centre of origin of invasive species, where the initial source populations can be found, should correspond to the centre of genetic diversity of the species. The underlying assumption is that introduced populations originate from a single source population through a genetic bottleneck and thus represent only a subset of the genetic diversity present in the source population (founder effect, Mayr 1954). Although reasonable in many cases, this assumption does not hold as a general rule. For instance, in the invaded areas, populations derived from admixture between several source populations may be as diverse as, or even more diverse than, the original source populations at the centre of origin (Kolbe *et al.* 2004; Facon *et al.* 2008). In plant pathogens, recent agricultural practices can also impact on the genetic diversity of invasive populations. For instance, in the Corn Smut pathogen *U. maydis* originating in Mexico, higher levels of diversity were found in more recent South-American and North-American populations, probably resulting from larger effective size of populations in these vast plantations (Munkacsy *et al.* 2008).

Overall, both single and multiple introductions have proved to lead to large-scale emergence with drastic or less tight bottlenecks depending on the species. Multiple introductions have been described in the wheat pathogens *Phaeosphaeria nodorum* and *M. graminicola* (Banke & McDonald 2005; Stukenbrock *et al.* 2006). By contrast, invasion of the European continent because of a single introduction has been suggested for the pathogen *Ceratocystis fimbriata* f. *platani* (Engelbrecht *et al.*

2004), and for the sugarcane fungal pathogen *Ustilago scitaminea*, for which a single lineage has been disseminated worldwide (Raboin *et al.* 2007). Thus, multiple introductions do not appear to be the rule for successful emerging diseases caused by aerial fungal pathogens, and other cases should be studied to further understand the role in pathogen emergence of demographic events such as bottlenecks and admixture.

Finally, another important aspect of fungal emergence that is frequently unknown is the relative importance of human transport vs. natural dispersal during range expansion. Most fungal crop pathogens produce wind-dispersed spores. While many emerging diseases of cultivated plants are known to result from anthropogenic introductions (Anderson *et al.* 2004), epidemiological studies have proposed the potential role of aerial long-distance dispersal events over intercontinental scale for these species (reviewed in Brown & Hovmøller 2002). Although natural vs. human modes of dispersal remain difficult to disentangle, the reconstruction of historical dissemination of pathogens is required to generate more precise hypotheses. Such diverse modes of dispersion may account for the highly contrasting introduction scenarios that have been described for fungal plant pathogens to date.

The black leaf streak disease (BLSD) of bananas (Rhodes 1964), caused by the Ascomycete *Mycosphaerella fijiensis*, is a remarkable model of a recent worldwide emerging crop disease, responsible for severe damage on banana leaves (Churchill 2011). The relative contributions of wind-dispersed spores issued from sexual reproduction and human-mediated vection of infected plant material to global range expansion are not known. Historical observations and hypotheses about the outbreaks and propagation of the disease around the world are available (reviewed in Jones 2000) but do not by themselves constitute reliable indications to track routes of introduction. While its host plant has been established across the tropics and subtropics for several centuries (Perrier *et al.* 2011), the emergence of *M. fijiensis* is recent (first mention in Fiji in 1963, Rhodes 1964), and the species centre of origin a priori unknown. The New Guinea-Solomon Islands have been proposed as the centre of origin, from which global invasion first started, based mainly on the dates of first mention in various countries (Stover 1978). Population genetic studies based on a few microsatellite and/or RFLP markers reported high levels of molecular diversity in Papua New Guinea (PNG) in comparison with other regions (Africa, Latin America and Pacific Islands), but also in the Philippines (Carlier *et al.* 1996). Thus, it was proposed that the actual centre of diversity might encompass the whole of South-East Asia (including PNG), in common with its host genus *Musa* (reviewed in Jones



2000). This hypothesis should be tested with *M. fijiensis* samples from other South-East Asian countries in order to further circumscribe and characterize the centre of *M. fijiensis* genetic diversity. In addition, the loss of diversity observed in all other continents has raised the hypothesis that introductions originated from South-East Asian populations with founder effects. Currently, neither the number and source of introductions, nor the intensity of bottlenecks is known.

In the present study, we studied the worldwide genetic diversity of *M. fijiensis*, in order to address the following questions: (i) Can we circumscribe the species centre of origin and decipher the initial steps of emergence? (ii) Can we characterize disease introduction events among different continents and detect the potential occurrence of bottlenecks and/or admixture? (iii) Do these results allow the underlying dispersal processes to be further characterized?

We based our analysis on a unique collection of samples collected in a variety of localities from 37 countries across the world. Various molecular analyses were performed on a set of 21 microsatellite markers and 8 nuclear sequence-based markers, and the combined results were aligned with historical data. Our study confirms a South-East Asian centre of origin of the pathogen and provides new insights into its initial emergence and global and continental spread. While the importance of founder events is confirmed throughout all continents, our results reveal a remarkable contrast between scenarios of introduction into the invaded African and American continents.

## Materials and methods

### Studied model

*Mycosphaerella fijiensis* is a heterothallic fungus displaying both an asexual and a regular sexual reproduction at almost all disease cycles (ascospores are produced after 30–40 days in favourable conditions, Jones 2000; Churchill 2011). As spores (conidia) produced asexually can disperse over scales of a few metres only, wind-dispersed ascospores are considered a more important source of inoculum for the spatial spread of the disease as they might be carried over several hundreds of kilometres (Jones 2000). Anthropogenic activities can also accelerate dispersion of the disease by the transport of infected plant material (planting or packing material). The first report of damage caused by *M. fijiensis* was in 1963 in Fiji (Rhodes 1964). Surveys following this first discovery argue that the disease would have been present widely in South-East Asia long before 1963, probably as a less severe form difficult to differentiate from the related species *Mycosphaerella musicola* (reviewed in

Jones 2000). The first characterization in Fiji actually occurred at the time of an unprecedentedly severe and damaging epidemic.

### Sample collection and DNA extraction

Samples were derived from a historical laboratory collection preserved in glycerol at  $-80^{\circ}\text{C}$ . Each individual corresponds to a strain derived from a single ascospore isolated from an infected banana leaf. DNA extraction from growing mycelium was carried out as described in Halkett *et al.* (2010).

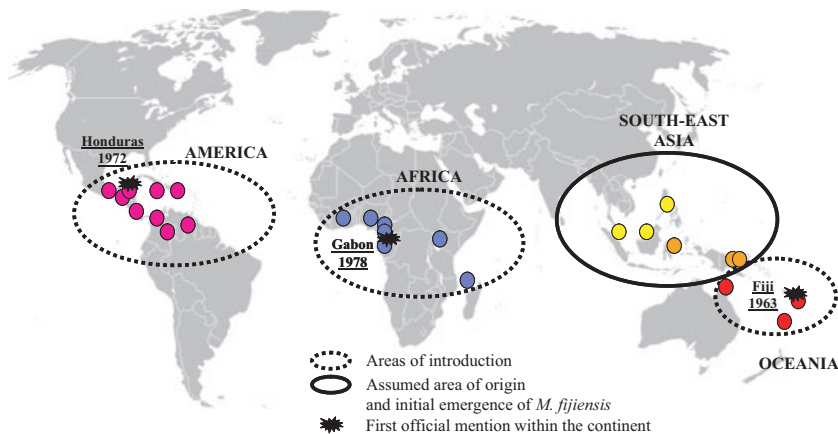
Samples were collected across the four large banana-producing continents: South-East Asia (including PNG), and three other continents: the Oceanian islands, Africa and America. In order to maximize the diversity of geographical localities considered, the 735 individuals included in this study were drawn from two types of sampling strategies: on the one hand, 25 distinct populations (629 individuals) from 22 different countries (Fig. 1) and, on the other, 106 additional individuals scattered throughout a variety of geographical localities where population samples were not available. Nineteen populations represented individuals collected simultaneously in the same banana field (hereafter referred to as 'local populations'). The six remaining populations, from Indonesia, Papua New Guinea, New Caledonia, Jamaica and Gabon, actually combine individuals originating from different localities (and will be referred to as 'composite populations'). Most banana hosts were susceptible clones belonging to the AAA and AAB genome groups. The global sampling design is presented in Table 1.

### Microsatellite genotyping

All individuals listed in Table 1 were genotyped using microsatellite markers. Twenty-one di-, tri- and tetra-nucleotide markers were retained: 9 were described in Neu *et al.* (1999) and Zapater *et al.* (2008), and the 12 other tetranucleotide markers were described in Robert *et al.* (2010). The markers were combined in three multiplex panels of seven markers each (Table S1) PCR amplification and genotyping were performed as described in Robert *et al.* (2010).

### Microsatellite diversity analysis

Gametic linkage equilibrium between all pairs of loci was tested in all populations using Fisher's exact tests implemented in GENEPOP<sub>4.0</sub> (Raymond & Rousset 1995; Rousset 2008). Composite populations were not included in analyses as population structure can artificially increase linkage disequilibrium. We used the false



**Fig. 1** Worldwide distribution of black leaf streak disease (BLSD) of bananas with the assumed area of origin and areas of introduction. Principal historical dates of first mention of the disease in different continents are indicated (according to Jones 2000). Sampling countries of our 25 *Mycosphaerella fijiensis* populations are marked by coloured circles.

discovery rate (FDR) procedure implemented in the R package QVALUE (Storey 2002) to control for multiple testing. Gene diversity  $H_E$  was estimated following Nei (1978), using  $F_{STAT_{2.9.3}}$  (Goudet 1995). To analyse allelic diversity while correcting for sample size differences, we used the rarefaction method implemented in HP-Rare (Kalinowski 2005) to compute allelic richness  $A_r$  and private allelic richness  $pA_r$ . These three diversity indices were computed at two levels: first for each of the 25 populations, and at the level of continents from the total genotyped data set. Differences between geographical areas for the three above parameters were then tested by performing a Wilcoxon signed rank test paired by loci in the software  $R_{2.13.1}$  (R Development Core Team 2011).

We used an AMOVA framework, as implemented in Arlequin<sub>3.11</sub> (Excoffier *et al.* 2005), to evaluate the degree of differentiation among continents and among populations within continents. Detailed pairwise  $F_{ST}$  (Weir & Cockerham 1984) between populations within each continent and  $F_{ST}$  between continents were also computed in  $F_{STAT}$ . The program Bottleneck (Cornuet & Luikart 1997) was used on the population data set to detect potential signals of recent bottlenecks, based on the detection of heterozygosity excess compared to heterozygosity expected under the mutation-drift equilibrium from the observed number of alleles. The underlying assumption is that the decrease in observed heterozygosity after a recent bottleneck is lower than the reduction in allelic diversity because of the extensive loss of rare alleles. Genetic distances among populations were estimated using Cavalli-Sforza chord distance and represented using a neighbour-joining tree, as implemented in POPULATIONS<sub>1.2.31</sub> software ([www.bioinformatics.org/~tryphon/populations/](http://www.bioinformatics.org/~tryphon/populations/)). Support values for each node of the tree were obtained with a bootstrap procedure of 1000 replications.

The most probable number of clusters supported by microsatellite data was estimated on the total

genotyped data set using the model-based procedure implemented in Structure<sub>2.3.3</sub> (Pritchard *et al.* 2000). Structure uses individual multilocus genotype data to cluster individuals into  $K$  groups while minimizing linkage disequilibrium between loci within groups and assigns individuals to these  $K$  clusters. We conducted a series of 10 independent runs for  $K$  varying from 1 to 12, allowing a model with admixture. Each run used  $10^6$  iterations after a burn-in of  $3.10^4$ . Markov Chain Monte Carlo (MCMC) convergence and consistency among the 10 replicated runs were checked for each value of  $K$ . The number of clusters that best explain our data was determined following the approach of Pritchard *et al.* (2000), based on the maximization of the logarithm likelihood of the data  $\ln \Pr(D|K)$ , and the approach of the Evanno's estimator  $\Delta K$  (Evanno *et al.* 2005), based on the second-order rate of change of  $\ln \Pr(D|K)$  when  $K$  increases. Clustering results were represented using DISTRICT<sub>1.1</sub> (Rosenberg 2004). To complement this Bayesian clustering approach, we used multivariate methods that make no assumptions regarding the data structure or population genetics model. A Discriminant analysis of principal components (DAPC, Jombart *et al.* 2010) was thus performed in the R package adegenet (Jombart 2008) to identify and describe clusters from the same data set. DAPC uses sequential K-means and model selection to infer genetic clusters: the Bayesian Information Criterion (BIC) allowed us to determine the optimal number of clusters. To describe the identified clusters, DAPC relies on data transformation using principal component analysis (PCA) as a prior step to DA (discriminant analysis), which maximizes the separation between groups. Further DAPC were also performed in order to investigate more precisely the potential source populations of the African and American samples: they are indeed more or less directly derived from South-East Asian and/or Oceanian populations as they are the most recently invaded areas.

**Table 1** Origin and sampling date of the 735 *Mycosphaerella fijiensis* individuals considered in the present study, including 25 populations (highlighted in grey) and scattered individuals. The whole sample has been genotyped with microsatellites, and individual contribution to nuclear intron sequencing is indicated

Geographic area	Country	Populations*	No. of individuals	Sampling date <sup>†</sup>	Sampling location <sup>†</sup>	Collector <sup>†</sup>	No. of individuals for intron-based markers
South-East Asia	China		6				3
	Taiwan		3				3
	Viet Nam		7				3
	Philippines	PHL	22	1993		X. Mourichon (CIRAD)	2
			2				2
	Malaysia	MAL1	13	1995	Borneo	R. Ploetz (University of Florida)	2
		MAL2	22	1995	Langkawi	H. Tezenas (CIRAD)	2
			1				1
	Indonesia	Ind	14	1995	Dai, Kao	I. Buddenhagen	3
			6				3
Oceania	Papua New Guinea	Png1	24	1989	several	B. Fullerton (PFR)	10
		Png2	25	2010	Kiriwo, Tarakbits	G. Kema (Wageningen UR), P. Kokoa (NAQIA)	–
	Australia	AUS	31	1996	Murray Island	E. Aitken (University of Queensland)	3
			1				–
	Fidji	FIJ	25	1996	Kosovou	C. Lavigne (CIRAD)	3
	New Caledonia	Ncl	22	1991	several	L. de Lapeyre (CIRAD)	3
	Futuna		4				3
	Vanuatu		1				1
	Samoa		3				3
	Tonga		3				3
America	Cook Islands		4				3
	Honduras	H	32	1999	Yojoa, Cortes	M. Rivera (FHIA)	1
		H-G	25	2009	La Lima, Cortes	M. Rivera (FHIA)	–
			9				2
	Colombia	COL	23	2009	La paila	T. Lescot (CIRAD)	–
			13				4
	Costa Rica	CRI	33	1999	Limon	G. Rivas (CATIE)	2
			2				2
	Mexico	MEX	10	2003	Yucatan	A. James (CICY)	1
			2				2
Africa	Venezuela	VEN	25	2009	Santa Barbara	T. Lescot (CIRAD)	–
			5				–
	Panama	PAN	32	1999	Finca Jazmin	E. Zaldaña	2
			1				1
	Ecuador		1				1
	Florida		3				3
	Cuba		7				3
	Jamaica	Jam	25	1995	several	R. Gonsalves (BBR)	4
	Dominican Rep.	RD	31	1999	Dajabon	T. Polanco (IDIAF)	3
	Uganda	UGA	34	1998		D. Vuylsteke (IITA)	3
Africa	Comoros	COM	34	2000		T. Lescot (CIRAD)	3
	Rwanda		1				1
	Tanzania		1				1
	Burundi		7				3
	Central African Rep.		1				1
	Congo		1				1
	Cameroon	CAM1	21	2001	Foret	C. Abadie (CIRAD)	2
		CAM2	24	2001	Kombore Bafou	C. Abadie (CIRAD)	2
			4				2

Table 1 Continued

Geographic area	Country	Populations*	No. of individuals	Sampling date <sup>†</sup>	Sampling location <sup>†</sup>	Collector <sup>†</sup>	No. of individuals for intron-based markers
	Gabon	<i>Gab</i>	24	1998	<i>several</i>	J. Guyot (CIRAD)	3
	Ghana		4				3
	Nigeria	<i>NGA</i>	24	1999	Onne	A. Tenkouano (IITA)	2
		1				1	
	Cote d'Ivoire	<i>CIV</i>	34	1999		C. Abadie (CIRAD)	3
		2				2	
Total			735				120

\*Abbreviations of local populations are in capital letters; composite populations are italicized.

<sup>†</sup>Indicated for populations only (where known) for clarity.

BBR, Banana Board; CATIE, Centro Agronómico Tropical de Investigación y Enseñanza; CICY, Centro de Investigación Científica de Yucatán; CIRAD, Centre de coopération internationale en recherche agronomique pour le développement; FHIA, Fundación hondureña de investigación agrícola; IDIAF, Instituto dominicano de investigaciones agropecuarias y forestales; IITA, International Institute of Tropical Agriculture; NAQIA, The National Agriculture Quarantine and Inspection Authority; PFR, Plant and Food Research.

After excluding African and American samples from the DAPC, we projected them as supplementary individuals onto the discriminant functions and examined their assignment to the different South-East Asia/Oceania clusters.

#### Development and characterization of sequence-based markers

The genome sequence of *M. fijiensis* (available on the JGI website: <http://genome.jgi-psf.org/Mycfi1/Mycfi1.home.html>) allowed us to develop eight sequence markers based on variable nuclear introns (Appendix S1). A total of 120 *M. fijiensis* individuals, selected from among both populations and scattered individuals, were sequenced at the French National Sequencing Centre (Genoscope). Sequences were aligned with the CLUSTALW algorithm and checked and corrected manually in BIOEDIT<sub>7.0</sub> (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>). Each marker was characterized by the haplotype diversity  $H_d$  and nucleotide diversity  $\pi$  calculated using DnaSP<sub>5</sub> (Librado & Rozas 2009; Table S2).

#### Analysis of sequence-based marker variation

Nucleotide diversity  $\pi$ , haplotype diversity  $H_d$  (Nei 1987) and the average number of nucleotide differences  $k$  (Tajima 1983) were computed in DnaSP<sub>5</sub> for each of the four continents, and differences between areas were tested across loci by performing a Wilcoxon signed rank test in the software R<sub>2.13.1</sub> (R Development Core Team 2011). Given the recent colonization history of the pathogen and its widespread sexual reproduc-

tion, the different loci of a sample were not expected to support the same genealogical histories; indeed, phylogenetic tree reconstructions proved inappropriate to analyse our sequence alignments. The SNAP Workbench java program (Price & Carbone 2005; Aylor *et al.* 2006) was thus used to examine the importance of phylogenetic conflict between sites along the multilocus alignment by drawing a compatibility matrix. An Incongruence Length Difference test (ILD, Farris *et al.* 1995) was performed in PAUP<sub>4.0</sub> (Swofford 2002) to test whether the level of incongruence observed between markers was greater than that expected by chance alone. Sequence alignments were collapsed into unique haplotypes for each marker separately using the module SNAP Map, and the minimum number of recombination events  $R_m$  (Hudson & Kaplan 1985) was inferred using the module RecPars in SNAP Workbench. To visualize the genetic connections between the different haplotypes, haplotype networks were constructed for each of the eight markers, using the median-joining method implemented in Network<sub>4.5.1.6</sub> (Fluxus Technology, Sudbury, UK).

Reticulate network approaches were then applied on multilocus data set to display the branching pattern and conflicting information supporting by all loci. The concatenated alignment of all markers was first collapsed into unique haplotypes in SNAP Map. A split network was computed by the method of neighbour-net in Splitstree4 (Huson & Bryant 2006) from a matrix of between-sequence genetic distances. The median-joining approach was also used in Network<sub>4.5.1.6</sub> (Fluxus Technology), which connects haplotypes directly from the patterns of substitutions in sequences alignment.

**Table 2** Genetic diversity indices estimated for both microsatellites and sequences data sets, at the levels of populations and continents (abbreviations as in Table 1)

Microsatellite markers variation									Sequence-based markers variation				
Total dataset (735 ind.)				Populations dataset (629 ind.)									
Geographic area (sample size)	H <sub>E</sub>	Ar	pAr	Population (sample size)	% of polymorphic loci	H <sub>E</sub>	Ar	Private Ar	Geographic area (sample size*)	No. of informative sites/total	π	Hd	k
S-E Asia (145)	0.65	8.3	4.6	Ind (14)	90 (19/21)	0.62	3.9	0.6	S-E Asia (23)	70/4198	0.0069	1	28.6
				MAL1 (13)	80 (17/21)	0.47	3.0	0.4					
				MAL2 (22)	71 (15/21)	0.36	2.3	0.21					
				PHL (22)	95 (20/21)	0.63	4.6	0.68					
				Png1 (24)	83 (15/18)	0.42	2.6	0.21					
				Png2 (25)	100 (21/21)	0.56	3.9	0.16					
Oceania (94)	0.33	2.6	0.1	FIJ (25)	62 (13/21)	0.19	1.7	0.05	Oceania (22)	25/4198	0.0028	0.96	11.7
				Ncl (22)	57 (12/21)	0.20	1.9	0					
				AUS (31)	81 (17/21)	0.36	2.6	0.06					
America (279)	0.35	2.6	0.17	COL (23)	65 (13/20)	0.31	2.2	0.03	America (28)	18/4198	0.0016	0.87	6.7
				CRI (33)	76 (16/21)	0.33	2.1	0.03					
				H (32)	71 (15/21)	0.37	2.2	0					
				H-G (25)	70 (14/20)	0.44	2.3	0.01					
				MEX (10)	71 (15/21)	0.43	2.2	0					
				PAN (32)	71 (15/21)	0.30	2.1	0					
				VEN (25)	55 (11/20)	0.31	2.1	0					
				Jam (25)	62 (13/21)	0.26	2.1	0.01					
				RD (31)	48 (10/21)	0.19	1.8	0.02					
				COM (34)	29 (6/21)	0.09	1.3	0					
Africa (217)	0.22	2.4	0.25	UGA (34)	81 (17/21)	0.31	2.3	0.06	Africa (29)	10/4198	0.0009	0.87	3.8
				CAM5 (21)	52 (11/21)	0.22	1.8	0.07					
				CAM6 (24)	57 (12/21)	0.18	1.7	0.01					
				CIV (34)	52 (11/21)	0.22	2.0	0					
				Gab (24)	57 (12/21)	0.23	1.7	0					
				NGA (24)	38 (8/21)	0.12	1.6	0					

H<sub>E</sub>, unbiased estimate of gene diversity (Nei 1978); Ar, allelic richness corrected for sample size; pAr, private allelic richness corrected for sample size (standardization to a sample size of 15 individuals for populations and 180 individuals for continents); π, nucleotide diversity (Nei 1987); Hd, haplotype diversity (Nei 1987); k, average number of nucleotide differences (Tajima 1983).

\*Excluding sequences with missing data not usable for calculations.

## Results

### *Microsatellite genetic diversity and global genetic structure*

Considering the 'local populations', no significant linkage disequilibrium between any pairs of loci was noted after FDR control. The measures of genetic diversity are displayed in Table 2. At continental level, Wilcoxon signed rank test indicated that significantly higher levels of variability were observed in South-East Asia (including PNG), in terms of either gene diversity (H<sub>E</sub>: 0.65), allelic richness (Ar: 8.3) or private allelic richness (pAr: 4.6). Africa tended to display the lowest amount of genetic diversity (H<sub>E</sub>: 0.22, Ar: 2.4) except for private richness (pAr: 0.25). Such tendencies were also verified

when considering the populations separately, in particular South-East Asian populations displayed higher diversity (H<sub>E</sub>: average  $0.51 \pm 0.11$ ; Ar: average  $3.4 \pm 0.9$ ; pAr: average  $0.4 \pm 0.2$ ) in comparison with populations in the rest of world (H<sub>E</sub>: average  $0.27 \pm 0.10$ ; Ar: average  $2.0 \pm 0.3$ ; pAr: average  $0.02 \pm 0.02$ ). Non-negligible differences were nevertheless noted between populations within areas. Within South-East Asia, the 'local population' of the Philippines exhibited the highest levels of diversity (H<sub>E</sub>: 0.63 and Ar: 4.6), with numerous private alleles (private Ar: 0.68), while the 'composite populations' from Indonesia and PNG showed no more variability than the other South-East Asian populations when estimated in this way. In America and Africa, the highest levels of variability were found in Honduras (H<sub>E</sub>: 0.44, Ar: 2.3) and in Uganda (H<sub>E</sub>: 0.31, Ar: 2.3),



**Table 3** Hierarchical analysis of molecular variance (AMOVA) of 25 populations of *Mycosphaerella fijiensis* distributed worldwide. In the global analysis, groups correspond to the four geographical areas defined in Table 1: South-East Asia, Oceania, America and Africa

Dataset	Number of populations	Percentage of variation populations			F-statistics
		Among groups	Among populations within groups	Within populations	
SE Asia	6	–	34.3	65.7	$F_{ST}$ 0.34***
Oceania	3	–	49.1	50.1	$F_{ST}$ 0.49***
America	9	–	24.7	75.3	$F_{ST}$ 0.25***
Africa	7	–	28.1	71.9	$F_{ST}$ 0.28***
					$F_{CT}$ 0.38***
Global	25	38.2	20.3	41.5	$F_{SC}$ 0.33***
					$F_{ST}$ 0.59***

\*\*\* $P < 0.01$ .

respectively. A trend towards a decrease in  $H_E$  with increasing geographic distance was noted in America, going from Honduras (0.44) to Venezuela (0.31).

The results of AMOVA analysis (Table 3) revealed that high levels of differentiation existed at all hierarchical levels ( $F_{ST}$  0.59,  $F_{SC}$  0.33,  $F_{CT}$  0.38; all values being highly significantly different from 0). While the greatest amount of variation was distributed within populations (41.5%), significantly large proportions of the variation are also attributable to differences between continents (38.2%), but also between populations within continents (20.3%). Such a variance distribution was also verified within the four continents. Table 4 showed consistently high values of  $F_{ST}$  within and between continents and revealed the more important differentiation of Africa from the other continents ( $F_{ST}$ : 0.52, 0.54 and 0.41 with America, Oceania and South-East Asia, respectively).

Significant signatures of bottlenecks on the distribution of allele frequencies were detected in several populations (Table S3). Notably, the most variable American populations of Honduras, Mexico and Costa Rica displayed highly significant signals. Significant results were also observed for other American populations and South-East Asian populations, and there were marginally significant results in some African populations. Although the number of microsatellites we used was too low to have sufficient power of analysis, especially in Africa, these results support the occurrence of widespread founder events having led to the establishment of these populations. The unrooted neighbour-joining tree constructed from chord distances based on microsatellites (Fig. 2) provides a graphical overview of the genetic relationships between populations. Clusters corresponded globally to geographical areas. Africa formed a single well-defined group that branched between Oceania and South-East Asia. American populations were also grouped together, close to Oceanian populations. South-East Asian populations were connected by long

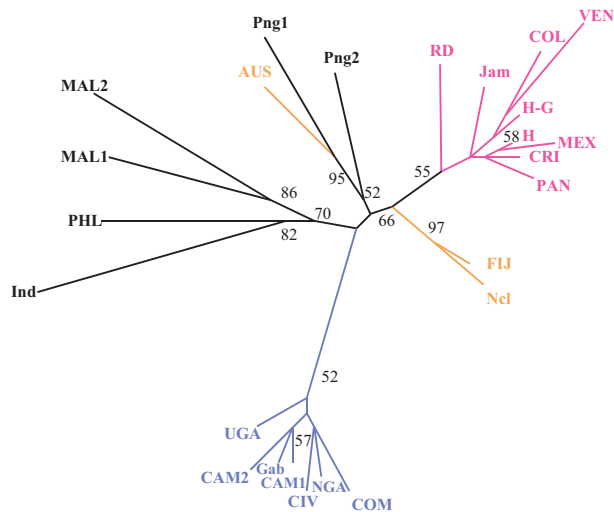
branches reflecting their high degree of genetic divergence compared to populations within other areas.

Based on the Bayesian clustering method Structure, the whole microsatellite data set was summarized efficiently with a cluster number ranging from  $K = 3$  to  $K = 6$ . MCMC convergence and consistency among runs were checked (no longer observed for  $K \geq 7$ ). According to Evanno's approach,  $K = 3$  was the best-fit number of genetic clusters, corresponding to the modal value of  $\Delta K$  (Fig. S1b). Following Pritchard *et al.* (2000), we nevertheless chose to examine consecutive solutions until  $K = 6$ , which led to a slightly smaller increase in the log likely hood of the data (Fig. S1a) but captured biologically meaningful substructure (Fig. 3). The three clusters for  $K = 3$  corresponded to South-East Asia, Africa and the grouping of America with Oceania. America and Oceania were then separated for  $K = 4$ . The number of clusters describing South-East Asia and Oceania then increased at higher values of  $K$ . For  $K = 6$  and  $K = 5$ , two truly distinct solutions were in fact highlighted among the 10 different runs, providing similarly appropriate explanations for the data. MCMC convergence and consistency among runs were verified within each solution. In the most frequent cases (80% of the runs), one very homogeneous cluster corresponded to the African samples, a second included all the American isolates, and the remaining cluster described a more complex clustering pattern in South-East Asia and Oceania. In 20% of the runs, a different clustering pattern was observed because of the presence of two clusters in America. These two clusters were distributed across all population samples, with the presence of individuals assigned preferentially to one cluster or the other and admixed individuals. Although these outputs must be interpreted carefully, these clustering patterns, when considered together with the results described above, raised the hypothesis of multiple introductions and admixture at the base of the

**Table 4** Pairwise  $F_{ST}$  (Weir & Cockerham 1984) estimated for all pairs of populations (abbreviations as in Table 1) and averaged within and between continents (values in bold)

Africa		America										Oceania					S-E Asia							
COM	UGA	CAM5	CAM6	CIV	GAB	NGA	COL	CR	H	H-G	MEX	PAN	VEN	JAM	RD	FJI	NCL	AUS	Ind	MAL1	MAL2	PHL	Png1	Png2
Africa	COM																							
	UGA	0.29																						
	CAM5	0.40	0.16																					
	CAM6	0.48	0.19	0.20																				
	CIV	0.34	0.25	0.21	0.22																			
America	GAB	0.41	0.09	0.07	0.17	0.30																		
	NGA	0.57	0.14	0.33	0.40	0.49	0.16																	
	COL	0.82	0.64	0.73	0.77	0.79	0.71	0.80																
	CR	0.69	0.52	0.57	0.65	0.66	0.58	0.69	0.33															
	H	0.68	0.51	0.55	0.62	0.63	0.56	0.67	0.30	0.02														
Oceania	H-G	0.73	0.55	0.60	0.67	0.68	0.60	0.72	0.19	0.13														
	MEX	0.77	0.54	0.61	0.69	0.71	0.61	0.76	0.41	0.17	0.12	0.23												
	PAN	0.69	0.50	0.56	0.63	0.66	0.56	0.67	0.30	0.07	0.10	0.20	0.23											
	VEN	0.81	0.62	0.70	0.75	0.77	0.70	0.80	0.34	0.37	0.32	0.21	0.47	0.42										
	JAM	0.76	0.59	0.66	0.73	0.73	0.66	0.77	0.42	0.13	0.08	0.24	0.26	0.21	0.48									
S-E Asia	RD	0.87	0.66	0.77	0.82	0.83	0.74	0.84	0.56	0.42	0.43	0.51	0.60	0.32	0.70	0.55								
	FJI	0.75	0.53	0.63	0.71	0.73	0.62	0.72	0.47	0.35	0.33	0.40	0.43	0.30	0.44	0.45	0.53							
	NCL	0.76	0.50	0.58	0.67	0.70	0.56	0.68	0.52	0.38	0.37	0.41	0.44	0.29	0.53	0.52	0.53	0.19	0.42					
	AUS	0.77	0.60	0.64	0.69	0.72	0.63	0.72	0.52	0.44	0.40	0.45	0.39	0.45	0.52	0.52	0.59	0.48	0.46					
	Ind	0.74	0.55	0.61	0.64	0.69	0.60	0.70	0.55	0.52	0.50	0.49	0.49	0.49	0.58	0.59	0.70	0.59	0.57	0.50				
	MAL1	0.78	0.57	0.64	0.69	0.73	0.63	0.72	0.52	0.50	0.46	0.47	0.49	0.49	0.51	0.56	0.71	0.54	0.53	0.49	0.39			
	MAL2	0.80	0.63	0.69	0.72	0.75	0.67	0.74	0.65	0.63	0.60	0.60	0.64	0.61	0.66	0.68	0.77	0.67	0.65	0.62	0.48	0.32		
	PHL	0.62	0.42	0.45	0.50	0.55	0.45	0.54	0.38	0.35	0.32	0.31	0.32	0.34	0.37	0.41	0.53	0.39	0.37	0.33	0.20	0.25	0.36	
	Png1	0.75	0.57	0.61	0.65	0.69	0.60	0.69	0.47	0.44	0.40	0.42	0.40	0.43	0.47	0.51	0.57	0.46	0.42	0.14	0.41	0.39	0.54	0.27
	Png2	0.60	0.36	0.43	0.51	0.54	0.42	0.53	0.38	0.28	0.24	0.26	0.24	0.26	0.24	0.35	0.34	0.46	0.23	0.18	0.23	0.37	0.33	0.50

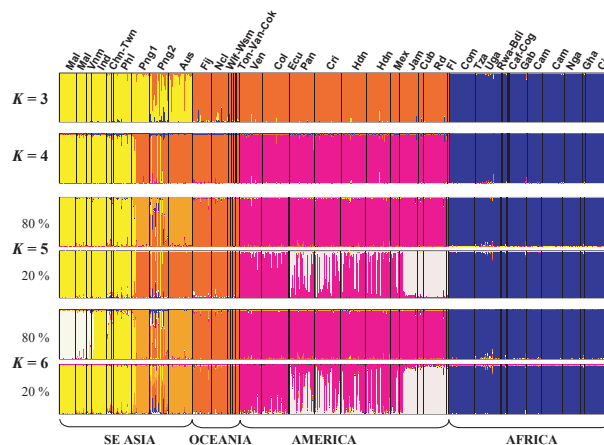
All values are significant according to permutation tests, except the two values in grey.



**Fig. 2** Unrooted neighbour-joining tree reconstruction based on Cavalli-Sforza microsatellite distances between the 25 populations (abbreviations correspond to those in Table 1). Numbers on the branches represent the percentage bootstrap support for values above 50%.

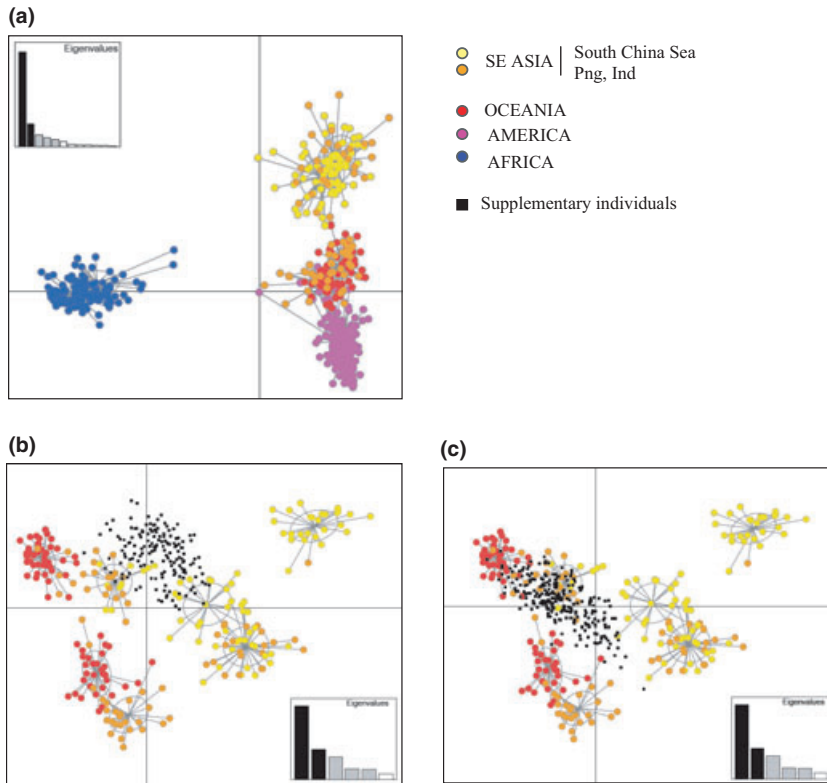
American invasion. The results also focus attention on the pattern of clustering displayed within PNG populations whatever the value of K: while South-East Asian samples were generally grouped within relatively homogeneous clusters, PNG populations appeared

much more structured. Especially in Png2—comprising numerous localities—a very mixed pattern representing all clusters could be observed within and/or between individuals. In DAPC analysis, the optimal number of clusters corresponds to the values after which the BIC increases or decreases by a negligible amount (Jombart *et al.* 2010). Here, values of K varying between K = 10 and K = 15 also represented a good summary of the data (Fig. S2a). A similar geographical structure was again revealed by DAPC analysis (Fig. 4a), together with the relative homogeneity and strong isolation of African individuals and the heterogeneity of America, which was described by a consistently higher number of clusters and appeared closely connected to Oceania and South-East Asia, especially PNG individuals. DAPC with supplementary individuals provide further information regarding the potential source populations of African and American introductions. Projection of supplementary African samples onto the South-East Asian/Oceanian DAPC (for K = 7, Fig. S2b) revealed their higher proximity with a cluster formed by Asian countries around the South China Sea (China, Taiwan, Philippines) and PNG, rather than with Oceania (Fig. 4b). Based on memberships probabilities of supplementary individuals, these analyses supported the assignment of 97% of African individuals to this same cluster (data not shown), consistently with the hypothesis of unique source population. Projection of American



**Fig. 3** Bayesian clustering of 735 *Mycosphaerella fijiensis* multilocus microsatellite haplotypes using Structure (model with admixture). Several outputs are presented, corresponding to the different values of K (from 3 to 6) that most efficiently summarize the data. A multimodality is observed for K = 5 and for K = 6, with two distinct solutions explored by Markov Chain Monte Carlo chains (percentage of estimations is indicated for each solution). Each individual is represented by a vertical line partitioned into K coloured segments that represent the probability that the haplotype belongs to each of K clusters. Black lines separate individuals from different countries and/or locality of origin, as indicated by abbreviations under the barplot. Bdi: Burundi, Com: Comores, Rwa: Rwanda, Tza: Tanzania, Uga: Uganda, Caf: Central Africa, Cam: Cameroon, Civ: Ivory Coast, Cog: Congo, Gab: Gabon, Gha: Ghana, Nga: Nigeria, Col: Colombia, Cr: Costa Rica, Hdn: Honduras, Mex: Mexico, Pan: Panama, Ven: Venezuela, Cub: Cuba, Jam: Jamaica, Rd: Dominican Republic, Fl: Florida, Cok: Cook islands, Fij: Fiji, Ncl: New Caledonia, Ton: Tonga, Van: Vanuatu, Wlf: Futuna, Wsm: Samoa, Aus: Australia, Chn: China, Twn: Taiwan, Vnm: Vietnam, Ind: Indonesia, Mal: Malaysia, Phl: Philippines, Png1 and Png2: Papua New Guinea (corresponding to the two composite populations).





**Fig. 4** Discriminant Analysis of Principal Components (DAPC, Jombart *et al.* 2010). (a) Scatterplot of DAPC (two first discriminant axis) performed on the global data set of 735 individuals, for  $K = 13$ . Individuals are represented by dots and clusters as inertia ellipses. Colours reflect geographical areas of samples. (b, c) Scatterplots of DAPC (two first discriminant axis) performed on South-East Asian and Oceanian samples, for  $K = 7$ . Supplementary African (b) and American (c) individuals are represented by black squares.

samples (Fig. 4c) showed clearly their genetic proximity with different clusters corresponding to Oceania on the one hand and South-East Asia on the other hand. This visual result was again supported by the memberships probabilities of American individuals with 65% assigned rather to South-East Asian clusters and 35% to Oceania, in agreement with the hypothesis of different sources populations.

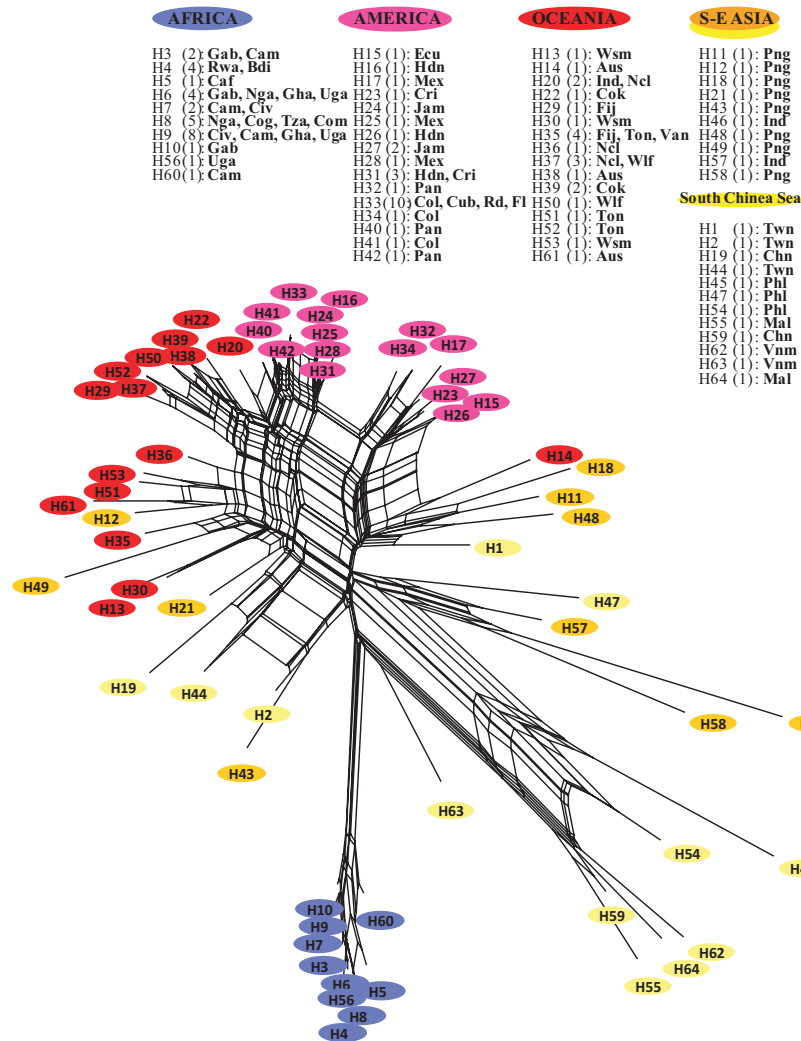
#### Sequence-based marker analysis

The results of diversity analysis for each continent are displayed in Table 2. A trend for higher nucleotide diversity (marginally significant), haplotype diversity ( $P$ -value = 0.004\*\*) and nucleotide divergence of sequences (marginally significant) was observed in South-East Asia as compared to all the other continents.

The compatibility analysis on the global sample of *Mycosphaerella fijiensis* revealed widespread phylogenetic conflicts among sites distributed mostly between loci (Fig. S3). A highly significant result of the ILD test ( $P$ -value < 0.01) confirmed that conflicts between markers reflected genuinely different genealogical history rather than artifactual noise. A few intralocus recombination events were also detected with the minimum number of within-locus recombination  $R_m$  varying from 0 to 3. Summaries of haplotype distributions and haplotype

networks for each marker separately are displayed in Table S4 and Fig. S4a, respectively. Despite the widespread incongruence of branching of individuals from one marker to another (not shown in detail here), some consistent information across the eight markers is deserving of note. Firstly, the higher diversity of South-East Asian haplotypes is seen here, compared to those of Africa, America and Oceania. Secondly, the few American and African haplotypes both displayed characteristic branching: the different African haplotypes always appeared directly linked and close to each other, while the different American haplotypes were systematically distant, as well as being separated by Asian haplotypes.

The split network computed from multilocus haplotypes (Fig. 5) displayed high levels of reticulation. The diversity and divergence among South-East Asian haplotypes was again seen, with PNG haplotypes scattered notably throughout the network. All African multilocus haplotypes were clustered together and split from the rest of the world with very well supported edges, supporting the common history of introduction of all African multilocus haplotypes. Indeed, it could be assumed that homoplasy because of convergent evolution was negligible given the low levels of polymorphism of our sequences. So, network reticulations were likely to reflect the history of introductions, reshaped by the widespread occurrence of pre- and post-intro-



**Fig. 5** Splits network for *Mycosphaerella fijiensis* haplotypes constructed by the method neighbour-net, based on combined data from all eight sequence-based markers. The different countries of origin of the 64 multilocus haplotypes are indicated in the legend (for abbreviations, see Fig. 3) and grouped by continents (within South-East Asia, two different groups are marked by colours: countries around the South China Sea and Papua New Guinea/Indonesia). Numbers in parentheses represent numbers of individuals sharing each haplotype. Each band of parallel edges indicates a split. The length of the edge representing a given split is an indication of its support.

duction recombination. Moreover, as noted in the legend to Fig. 5, some identical multilocus haplotypes could be shared by very distant individuals within Africa: for example, the African haplotype H9 was found from Uganda to the Ivory Coast. In other words, all our samples across Africa could be derived from a single introduction event characterized by a bottleneck that introduced only a few different haplotypes from a unique source population. Consistently with DAPC analyses, this African cluster was connected to haplotypes from South China Sea countries (including China, Taiwan, Vietnam, Malaysia, Philippines). American haplotypes were also grouped together, but they appeared separated into two groups by an extensive split in the network, reflecting notably widespread

conflicts among the genealogical histories supported by different recombination blocks. American haplotypes were strongly interconnected with South-East Asian haplotypes on the one side and Oceanian on the other side. Moreover, from the legend of Fig. 5, it could be noted that geographically distant populations frequently shared close multilocus haplotypes (e.g. H17 in Mexico and H34 in Colombia), while individuals from the same locality could display very divergent haplotypes (e.g. H34 and H41 in Colombia). Taken as a whole, these results strongly suggest that American multilocus haplotypes could be derived from a mix between South-East Asian and Oceanian haplotypes and that recombinant haplotypes have spread across the continent, consistent with the hypothesis of admixture raised by microsatel-

lites analyses. The same interpretation was supported by the multilocus median-joining network (Fig. S4b). With this method of reconstruction, a single long branch connected the African cluster to the rest of the world, while the American cluster appeared in a central position, being linked by multiple connections to South-East Asian haplotypes on the one side and Oceanian haplotypes on the other.

## Discussion

### *Centre of diversity of *Mycosphaerella fijiensis* and initial emergence of the disease*

Studying *Mycosphaerella fijiensis* worldwide population structure confirmed our previous hypothesis (Carlier *et al.* 1996) that the whole of South-East Asia could be the centre of diversity of *M. fijiensis*. This encompasses at least Malaysia, Philippines, Indonesia and PNG, but excluding the Oceanian islands (Fiji and New Caledonia). Moreover, the high genetic divergence of sequence markers between South-East Asian localities combined with a large number of private haplotypes within each locality, and high private allelic richness at the continent scale, support the hypothesis that the disease has long been present in this area.

It is worth noting that South-East Asia is the centre of origin of the host plant, displaying a diversity of wild banana and plantain species, and defining the area of banana and plantain domestication that began several thousand years ago (Perrier *et al.* 2011). A coevolution between *M. fijiensis* and its host could have occurred during the process of domestication of *Musa*, implying also that *M. fijiensis* could have been disseminated locally with domesticated banana. Thus, on the scale of the South-East Asian area, BLSD could represent a supplementary case of fungal emergence resulting from host-tracking. The question now is whether the pathogen originated over the whole of South-East Asia or within a narrower region of the continent? In support of the latter, New Guinea and the Solomon Islands were historically suspected to be the most likely native area of *M. fijiensis* Stover (1978). Stover reported that during the last century infected leaves or planting material could have been moved frequently during inter-island trade from New Guinea and the Solomon Islands to the rest of South-East Asia and to the Pacific islands, creating ideal conditions for pathogen dissemination. Another argument is that PNG is considered an important primary centre of banana domestication (Perrier *et al.* 2011). The results obtained from Structure and sequence-based networks, which show that PNG individuals are distributed in all possible clusters, support this hypothesis. However, the Structure output could

also result artificially from a random assignment of individuals to the different clusters, because of the high genetic structure within this composite population. To properly localize the *M. fijiensis* centre of origin, and to confirm the suspected host-tracking scenario within South-East Asia, pathogen populations from both domesticated and wild bananas should be compared. History and distribution of bananas in South-East Asia should be further investigated.

### *Deciphering global introduction events: contrast between Africa and America*

Significant reductions in genetic variability characterized all invaded areas. However, genetic structure differed among invaded continents. Our study confirms that Oceanian populations are related closely to those of PNG and may therefore have been introduced from this source. While the present sampling does not permit the number of introductions and the dissemination pathways within this island system to be inferred, a more important focus can be drawn in Africa and America.

One remarkable result of this study is the pattern of genetic diversity revealed in Africa, which forms an isolated and relatively homogenous cluster supported by microsatellite and sequence-based marker analyses. Different introduction events in Africa were already suspected from historical information to explain the rapid dissemination of the disease to both the East and West sides of the continent (Jones 2000): a first event either in central Africa or in Gabon, and another event near Tanzania. The present study rather strongly suggests a single introduction event characterized by a bottleneck, coming from a single South-East Asian source population around South China Sea. The surprisingly high number of private microsatellite alleles and private sequence haplotypes in Africa, added to the high differentiation of the continent from the rest of the world, might indicate a more ancient introduction event compared to the other invaded continents, despite the fact that this is neither in line with historical data nor with the low genetic diversity of African populations. A more plausible explanation may be that populations close to the real source population of the African introduction may not have been sampled. Following this introduction, stochastic founder events without any geographical logic would then have occurred across the continent, modifying allelic frequencies to generate the extensive differentiation observed between populations.

The American distribution of *M. fijiensis* genetic variability contrasts strongly with that seen in Africa. According to both microsatellite and sequence analyses, American populations could be derived from admixture between different genetically differentiated sources

from South-East Asia and Oceania. Admixture could actually have occurred either through simultaneous multiple introductions in the same place, or through multiple introductions at different times and places. The former would be more consistent with the distribution of microsatellite and sequence variability across the continent, and especially the trend towards a decrease in microsatellite genetic variability from Honduras to Venezuela. The high levels of variability found in Honduras populations suggest that they were established first, followed by successive bottlenecks during progression towards the South. The distribution of sequence haplotypes also supports a scenario of admixture within a founding population, followed by spread of the recombined genotypes across the continent. The complementarity of our two types of markers appears obvious here, as sequence markers preserve information about the origin of alleles, allowing us to keep track of ancestral shared polymorphism longer than hypervariable markers, which describe for their part the relationships between recently diverged populations. It is worth noting that this scenario of admixture appears strikingly consistent with historical reports of the first detection of *M. fijiensis* in the American continent. The pathogen was identified officially for the first time in Honduras in 1972, in a farm containing a large collection of bananas and plantains, but would have been present early in the 1960s (Stover & Dickson 1976). A population analysed in this study collected from this same farm (H-G) displayed the highest levels of genetic diversity in America. Interestingly, the banana collection was gathered in the 1960s from numerous locations throughout South-East Asia, especially Indonesia, Malaysia, the Philippines and PNG (Rosales *et al.* 1999), thus providing suitable conditions for a mix of differentiated South-East Asian individuals. However, our results suggest that an Oceanian source of *M. fijiensis* American populations must also be suspected, for example, the important trade route from the Pacific islands to America via Hawaii (Jones 2000) in the 1960s. Finally, in suggesting an admixture event at the introduction into a continent in a sexually reproducing fungus, our results appear unprecedented among previously described cases of plant pathogen invasions at the worldwide scale. Admixture in invasive fungal plant pathogens has in fact been demonstrated only rarely, for example, on the asexual sapstain fungus *Ophiostoma ips* (Zhou *et al.* 2007) or in *Phytophthora infestans* between clonal lineages at the country scale (Montarry *et al.* 2010). Even if admixture is more generally expected to frequently occur during invasion processes (Estoup & Guillemaud 2010), only a few studies of contemporary biological invasions have properly demonstrated its occurrence in invaded areas. These were often based on

both genetic and phenotypic trait diversity analysis and/or on a more exhaustive sampling of the native and invaded areas (Kolbe *et al.* 2004, 2007; Lavergne & Molofsky 2007; Facon *et al.* 2008; Keller & Taylor 2010). In the case of *M. fijiensis*, a more thorough sampling in South-East Asia and Oceania is first required to better identify the different source populations of the American invasion. Moreover, the available genome sequence of *M. fijiensis* appears as a valuable resource for further researches using molecular markers.

#### *Modes of dissemination across invaded continents*

The level of genetic differentiation between continents observed in *M. fijiensis* is much higher in comparison to other related crop pathogens (with regular sexual reproduction) such as the wheat pathogens *Mycosphaerella graminicola* and *Phaeosphaeria nodorum* and the apple pathogen *Venturia inaequalis*, but more similar to the barley pathogen *Rhynchosporium secalis* (Zaffarano *et al.* 2006, 2009). As for *R. secalis*, the global spread of *M. fijiensis* is very recent (a few decades only) in comparison with other pathogens, which followed the dissemination of their hosts thousands of years ago. Population structures established during pathogen spread are consequently still detectable for *M. fijiensis* and *R. secalis*. Because ascomycete spores are not very resistant to UV radiation (Parnell *et al.* 1998), it is generally thought that exchange of infected plant material is the most probable mode of dissemination between continents for all the above pathogens (in contrast with rusts, whose spores are UV resistant, Hovmøller *et al.* 2008). Here, genetic as well as historical data are consistent with human-mediated introductions into Africa and America. Such a high differentiation between invaded and native continents nevertheless suggests that infected plant material transport is less frequent in *M. fijiensis* than in other species (e.g. *Mycosphaerella graminicola* and *P. nodorum* and the apple pathogen *V. inaequalis*). One reason why human-mediated transport is limited in *M. fijiensis* may be that banana fruits—the main target of commercial exchange—are not infected by the fungus, unlike the cases of wheat or apple fruits. Furthermore, commercial exchange of plant material now relies mostly on the use of healthy vitroplants, with traditional banana multiplication through vegetative reproduction being of only minor importance.

A high level of genetic differentiation in *M. fijiensis* was also detected within continents consistently with the results of Rivas *et al.* (2004). This was a priori not expected given its abilities for wind dispersal, its widespread recombination, its perennial host crop and the homogeneity of banana agrosystems across regions. More localized studies have described absent or weak population structure in banana-producing areas of



several hundred kilometres wide, suggesting gradual spread of the disease through natural dispersal of spores (Halkett *et al.* 2010). However, high levels of genetic differentiation among populations at a larger scale could result from founder effects through either long-distance dispersal of spores or movement of infected plant materials. Landscape genetic studies from two-dimensional samplings may be useful to disentangle the relative importance of these different dispersal modes and help further understand the spread of plant pathogenic fungi (Manel *et al.* 2003; Rieux *et al.* 2011).

### Conclusion and perspectives

Overall, an original global panorama of sources, pathways and demographical characteristics of the introduction of a fungal plant pathogen can be drawn from this study, although the very first step of emergence of the pathogen remains to be further deciphered. While the occurrence of bottlenecks seems to be a general rule, this is not the case for multiple introductions. An unprecedented scenario of global spread was indeed revealed in *M. fijiensis*, with the occurrence of both single and multiple successful introductions. In particular, a single introduction event from a unique source in Africa and admixture events in America have both led to rapid and devastating invasions of these two continents. The effective contribution of multiple introductions and admixture to the success of emergence is especially deserving of further examination. Finally, our study calls for more precise investigation focused on both agrosystem history and pathogen aggressiveness traits, to dovetail with the growing number of studies assessing whether invasions result from the introduction of pre-adapted genotypes or are the consequence of post-introduction evolutionary shifts (Facon *et al.* 2006).

### Acknowledgements

We are grateful to Thierry Lescot (CIRAD, France), Mauricio Rivera (FHIA, Honduras), Gert Kema (Wageningen UR, Netherlands) and Pere Kokoa (NAQIA, Papua New Guinea) for collecting and sending new samples of *M. fijiensis*; Luc Pignolet (CIRAD, France) for help in isolation of fungal strains and DNA extractions; Arnaud Estoup (INRA, France), Benoit Facon (INRA, France), Elizabeth Fournier (INRA, France), Johan Michaux (FNRS, France), Adrien Rieux (CIRAD, France) and Renaud Vitalis (CNRS, France) for helpful discussions; Helen Rothnie for her attentive reading and careful English language checking. Data used in this work were produced through the molecular genetic analysis technical facilities of the IFR119 with the help of Erick Desmarais and Frédérique Cerqueira for microsatellite scoring, and through the Genoscope for DNA sequencing. This work was supported by the Région Languedoc Roussillon and the Agence Nationale de la Recherche, pro-

jects Emerfundis (ANR 07-BDIV-003) and EMILE (ANR 09-BLAN-0145-01).

### References

- Anderson PK, Cunningham AA, Patel NG *et al.* (2004) Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution*, **19**, 535–544.
- Aylor DL, Price EW, Carbone I (2006) SNAP: Combine and Map modules for multilocus population genetic analysis. *Bioinformatics*, **22**, 1399–1401.
- Banke S, McDonald BA (2005) Migration patterns among global populations of the pathogenic fungus *Mycosphaerella graminicola*. *Molecular Ecology*, **14**, 1881–1896.
- Barres B, Halkett F, Dutech C *et al.* (2008) Genetic structure of the poplar rust fungus *Melampsora larici-populina*: evidence for isolation by distance in Europe and recent founder effects overseas. *Infection Genetics and Evolution*, **8**, 577–587.
- Brewer MT, Milgroom MG (2010) Phylogeography and population structure of the grape powdery mildew fungus, *Erysiphe necator*, from diverse *Vitis* species. *BMC Evolutionary Biology*, **10**, 268.
- Brown JKM, Hovmöller MS (2002) Epidemiology – aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science*, **297**, 537–541.
- Carlier J, Lebrun MH, Zapater MF, Dubois C, Mourichon X (1996) Genetic structure of the global population of banana black leaf streak fungus, *Mycosphaerella fijiensis*. *Molecular Ecology*, **5**, 499–510.
- Churchill ACL (2011) *Mycosphaerella fijiensis*, the black leaf streak pathogen of banana: progress towards understanding pathogen biology and detection, disease development, and the challenges of control. *Molecular Plant Pathology*, **12**, 307–328.
- Cornuet JM, Luikart G (1997) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.
- Couch BC, Fudal I, Lebrun MH *et al.* (2005) Origins of host-specific populations of the blast pathogen *Magnaporthe oryzae* in crop domestication with subsequent expansion of pandemic clones on rice and weeds of rice. *Genetics*, **170**, 613–630.
- Desprez-Loustau ML, Robin C, Buee M *et al.* (2007) The fungal dimension of biological invasions. *Trends in Ecology & Evolution*, **22**, 472–480.
- Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, **17**, 431–449.
- Engelbrecht CJB, Harrington TC, Steimel J, Capretti P (2004) Genetic variation in eastern North American and putatively introduced populations of *Ceratocystis fimbriata* f. platani. *Molecular Ecology*, **13**, 2995–3005.
- Estoup A, Guillemaud T (2010) Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology*, **19**, 4113–4130.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.
- Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, **1**, 47–50.

- Facon B, Genton BJ, Shykoff J *et al.* (2006) A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology & Evolution*, **21**, 130–135.
- Facon B, Pointier JP, Jarne P, Sarda V, David P (2008) High genetic variance in life-history strategies within invasive populations by way of multiple introductions. *Current Biology*, **18**, 363–367.
- Farris JS, Kallersjo M, Kluge AG, Bult C (1995) Constructing a significance test for incongruence. *Systematic Biology*, **44**, 570–572.
- Frankham R (2005) Invasion biology – resolving the genetic paradox in invasive species. *Heredity*, **94**, 385–385.
- Gladieux P, Zhang XG, Afoufa-Bastien D *et al.* (2008) On the origin and spread of the scab disease of apple: out of Central Asia. *PLoS ONE*, **3**, e1455.
- Gladieux P, Zhang XG, Roldan-Ruiz I *et al.* (2010) Evolution of the population structure of *Venturia inaequalis*, the apple scab fungus, associated with the domestication of its host. *Molecular Ecology*, **19**, 658–674.
- Gomez-Alpizar L, Carbone I, Ristaino JB (2007) An Andean origin of *Phytophthora infestans* inferred from mitochondrial and nuclear gene genealogies. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 3306–3311.
- Goudet J (1995) FSTAT (Version 1.2): a computer program to calculate F-statistics. *Journal of Heredity*, **86**, 485–486.
- Halkett F, Coste D, Platero GGR *et al.* (2010) Genetic discontinuities and disequilibria in recently established populations of the plant pathogenic fungus *Mycosphaerella fijiensis*. *Molecular Ecology*, **19**, 3909–3923.
- Hovmøller M, Yahyaoui AH, Milus EA, Justesen AF (2008) Aggressive strains of the wheat yellow rust fungus spread world-wide. *Phytopathology*, **98**, S69–S69.
- Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics*, **111**, 147–164.
- Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, **23**, 254–267.
- Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.
- Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, **11**, 94.
- Jones DR (2000) Introduction to banana, abacá and enset. In: *Diseases of Banana, Abacá and Enset* (ed. Jones DR), pp. 37–79. CABI Publishing, Wallingford, UK.
- Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, **5**, 187–189.
- Keller SR, Taylor DR (2010) Genomic admixture increases fitness during a biological invasion. *Journal of Evolutionary Biology*, **23**, 1720–1731.
- Kolbe JJ, Glor RE, Schettino LRG *et al.* (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, **431**, 177–181.
- Kolbe JJ, Glor RE, Schettino LR *et al.* (2007) Multiple sources, admixture, and genetic variation in introduced Anolis lizard populations. *Conservation Biology*, **21**, 1612–1625.
- Lavergne S, Molofsky J (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 3883–3888.
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.
- Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, **18**, 189–197.
- Mayr E (1954) Change of genetic environment and evolution. In: *Evolution as a Process* (eds Huxley J, Hardy AC and Ford EB), pp. 157–180. Allen & Unwin, London.
- Montarry J, Andrivon D, Glais I *et al.* (2010) Microsatellite markers reveal two admixed genetic groups and an ongoing displacement within the French population of the invasive plant pathogen *Phytophthora infestans*. *Molecular Ecology*, **19**, 1965–1977.
- Munkacsı AB, Stoxen S, May G (2008) *Ustilago maydis* populations tracked maize through domestication and cultivation in the Americas. *Proceedings of the Royal Society B-Biological Sciences*, **275**, 1037–1046.
- Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.
- Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Neu C, Kaemmer D, Kahl G, Fischer D, Weising K (1999) Polymorphic microsatellite markers for the banana pathogen *Mycosphaerella fijiensis*. *Molecular Ecology*, **8**, 523–525.
- Novak SJ (2007) The role of evolution in the invasion process. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 3671–3672.
- Parnell M, Burt PJA, Wilson K (1998) The influence of exposure to ultraviolet radiation in simulated sunlight on ascospores causing Black Sigatoka disease of banana and plantain. *International Journal of Biometeorology*, **42**, 22–27.
- Perrier X, De Langhe E, Donohue M *et al.* (2011) Multidisciplinary perspectives on banana (*Musa* spp.) domestication. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 11311–11318.
- Price EW, Carbone I (2005) SNAP: workbench management tool for evolutionary population genetic analysis. *Bioinformatics*, **21**, 402–404.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- R Development Core Team (2011) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Raboin LM, Selvi A, Oliveira KM *et al.* (2007) Evidence for the dispersal of a unique lineage from Asia to America and Africa in the sugarcane fungal pathogen *Ustilago scitaminea*. *Fungal Genetics and Biology*, **44**, 64–76.
- Raymond M, Rousset F (1995) GENEPOP (Version-1.2) – population-genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.
- Rhodes PL (1964) A new banana disease in Fiji. *Commonwealth Phytopathological News*, **10**, 38–41.
- Rieux A, Halkett F, de Bellaire LD *et al.* (2011) Inferences on pathogenic fungus population structures from microsatellite

- data: new insights from spatial genetics approaches. *Molecular Ecology*, **20**, 1661–1674.
- Rivas GG, Zapater MF, Abadie C, Carlier J (2004) Founder effects and stochastic dispersal at the continental scale of the fungal pathogen of bananas *Mycosphaerella fijiensis*. *Molecular Ecology*, **13**, 471–482.
- Robert S, Rieux A, Argout X, Carlier J, Zapater MF (2010) Optimized genotyping with microsatellite markers in the fungal banana pathogen *Mycosphaerella fijiensis* (Mycosphaerellaceae). *American Journal of Botany*, **97**, E130–E132.
- Rosales FE, Tripon SC, Cerna J (eds) (1999) *A Tribute to the Work of Paul H. Allen: A Catalogue of Wild and Cultivated Bananas*. INIBAP/CIID/EARTH, Montpellier, France.
- Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.
- Rousset F (2008) GENEPOP '007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.
- Storey JD (2002) A direct approach to false discovery rates. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, **64**, 479–498.
- Stover RH (1978) Distribution and probable origin of *Mycosphaerella fijiensis* in South East Asia. *Tropical Agriculture, Trinidad*, **55**, 65–68.
- Stover RH, Dickson JD (1976) Banana leaf spot caused by *Mycosphaerella musicola* and *M. fijiensis* var. *difformis*: a comparison of the first Central American epidemics. *FAO Plant Protection Bulletin*, **24**, 36–42.
- Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology*, **46**, 75–100.
- Stukenbrock EH, Banke S, McDonald BA (2006) Global migration patterns in the fungal wheat pathogen *Phaeosphaeria nodorum*. *Molecular Ecology*, **15**, 2895–2904.
- Stukenbrock EH, Banke S, Javan-Nikkhah M, McDonald BA (2007) Origin and domestication of the fungal wheat pathogen *Mycosphaerella graminicola* via sympatric speciation. *Molecular Biology and Evolution*, **24**, 398–411.
- Swofford DL (2002) *PAUP\*. Phylogenetic Analysis Using Parsimony (\*and Other Methods)*. Sinauer Associates, Sunderland, Massachusetts.
- Tajima F (1983) Evolutionary relationship of DNA sequences in finite populations. *Genetics*, **105**, 437–460.
- Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
- Zaffarano PL, McDonald BA, Zala M, Linde CC (2006) Global hierarchical gene diversity analysis suggests the Fertile Crescent is not the center of origin of the barley scald pathogen *Rhynchosporium secalis*. *Phytopathology*, **96**, 941–950.
- Zaffarano PL, McDonald BA, Linde CC (2009) Phylogeographical analyses reveal global migration patterns of the barley scald pathogen *Rhynchosporium secalis*. *Molecular Ecology*, **18**, 279–293.
- Zapater MF, Duchemin M, Dussart JF *et al.* (2008) Microsatellite markers for the fungal banana pathogens *Mycosphaerella fijiensis*, *Mycosphaerella musicola* and *Mycosphaerella eumusae*. *Molecular Ecology Resources*, **8**, 1121–1125.
- Zhou XD, Burgess TI, Beer ZW *et al.* (2007) High intercontinental migration rates and population admixture in the sapstain fungus *Ophiostoma ips*. *Molecular Ecology*, **16**, 89–99.

This study was conducted by S.R. during her PhD thesis in the team BECφ of the research unit BGPI. The main objective of this PhD thesis is to infer global invasion patterns of *Mycosphaerella fijiensis* using different phylogeographical and population genetics approaches, and to study global variability of life history traits in this model. M-F.Z., C.A. and J.C. have a common research program dealing mainly with the dynamics and evolution of populations of the fungus *Mycosphaerella fijiensis* pathogenic on banana. V.R. adapts population genetic methods and develops models to study the evolution of plant pathogenic fungi.

## Data accessibility

Microsatellite and sequence data deposited in the Dryad repository: doi: 10.5061/dryad.kq74g8p7.

## Supporting information

Additional supporting information may be found in the online version of this article.

**Fig. S1** Structure analyses.

**Fig. S2** Inference of the number of clusters in the DAPC analyses performed (Discriminant analysis of principal components (DAPC, Jombart *et al.* 2010) performed (a) on the total global sample of 735 individuals (b) on South-East Asian and Oceanian samples.

**Fig. S3** Pairwise compatibility matrix for all informative sites from all eight sequence-based markers (excluding indels and infinite site violations), generated in SNAP Workbench on the global dataset of 120 individuals.

**Fig. S4** Median joining network constructed from the global sample of 120 *M. fijiensis* individuals (a) for the eight markers separately (b) for multilocus haplotypes based on combined data from all eight sequence-based markers.

**Table S1** Reference, multiplexing information and genetic variability of the 21 microsatellite markers used in this study

**Table S2** Molecular characterization and genetic variability of the eight intron-based nuclear markers developed for this study

**Table S3** Results of several tests run on the 25 *M. fijiensis* populations in the software Bottleneck, in order to test the existence of an heterozygosity excess observed in the samples (calculated as Nei's gene diversity, Nei 1987) compared to the heterozygosity expected under mutation-drift equilibrium and termed the allelic richness

**Table S4** Haplotype distribution of the global sample of *M. fijiensis* for the eight sequence-based markers

## Appendix S1 Development of nuclear intron-based markers

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

## **Supporting information**

### **Appendix 1: Development of nuclear intron-based markers.**

Two sources of data helped us to find neutral polymorphisms at the intraspecific level in nuclear DNA: first, the classical sequence-based markers used in published phylogenetic studies in Ascomycetes—specifically in the *Mycosphaerella* genus (e.g., Arzanlou *et al.* 2008); and second, the recently established database FunyBASE—a repository of ortholog families of genes well-suited for phylogenetic studies in Ascomycetes (Marthey *et al.* 2008; <http://genome.jouy.inra.fr/funybase/>). The genome sequence of *M. fijiensis* (available on the JGI website: <http://genome.jgi-psf.org/Mycfi1/Mycfi1.home.html>) allowed us to target the most variable nuclear introns flanked by conserved exons. Fourteen primer pairs were thus designed with PRIMER 3 (<http://frodo.wi.mit.edu/primer3/>, Rozen & Skaletsky 2000) and tested for PCR amplification on a small global sample *M. fijiensis*, to retain 12 sequences.

Their polymorphism was then tested on an initial global sample of 28 *M. fijiensis*, 2 *M. musicola* and 2 *M. eumusae* individuals. PCR reactions were performed in a PTC-100 thermocycler under the following optimized conditions: an initial step of 7min at 94°C, 35 cycles including 1min at 94°C, 1min at 55°C or 55.5°C depending on the primers used (Table S2), and 3min at 72°C. The 30µL reaction mixture included 6µL DNA (5ng/µL), 3µL 1X reaction buffer (Eurobio), 3µL dNTP, 1.2µL MgCl<sub>2</sub>, 0.12µL *Taq* DNA-polymerase (Eurobio), 0.03µL of each primer at a concentration of 0.1µM, and ddH<sub>2</sub>O up to a final volume of 30µL. PCR products were sequenced from both sides at the French National Sequencing Centre (Genoscope). The sequences were aligned with the ClustalW algorithm, and checked and corrected manually in BIOEDIT<sub>7.0</sub> (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>). Eight



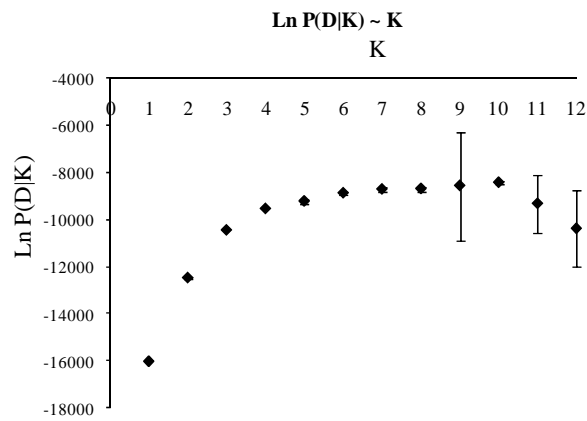
intron-based markers were finally retained. An additional sample was then sequenced with these 8 markers to obtain a total sequenced sample of 120 *M. fijiensis* individuals.

### *References*

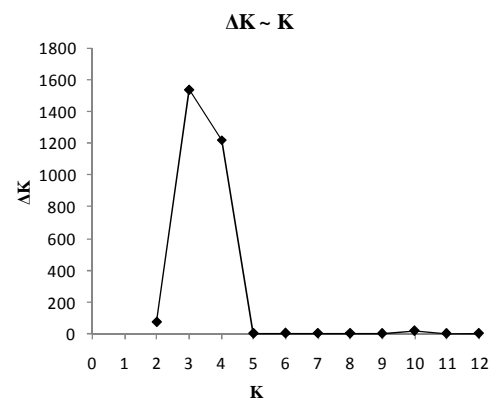
- Arzanlou M, Groenewald JZ, Fullerton RA, *et al.* (2008) Multiple gene genealogies and phenotypic characters differentiate several novel species of *Mycosphaerella* and related anamorphs on banana. *Persoonia* **20**, 19-37.
- Marthey S, Aguilera G, Rodolphe F, *et al.* (2008) FUNYBASE: a FUNgal phYlogenomic dataBASE. *Bmc Bioinformatics* **9**.
- Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**, 365-386.

**Figure S1: STRUCTURE analyses.** (a) Mean ( $\pm$  SD) log probabilities of the data  $\text{LnP}(D|K)$  over 10 STRUCTURE replicated runs plotted as a function of the putative number of clusters  $K$ . (b) Values of  $\Delta K$  calculated following Evanno *et al.* (2005) plotted as a function of the putative number of clusters  $K$ .

(a)

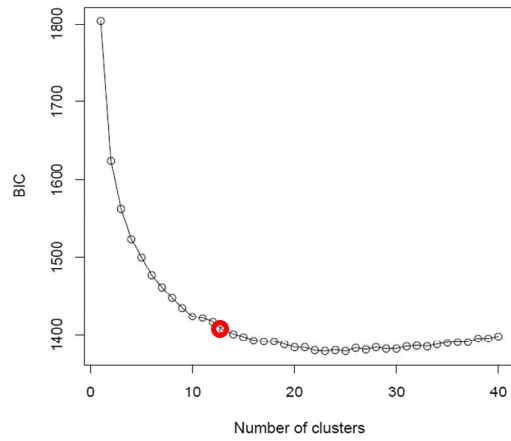


(b)

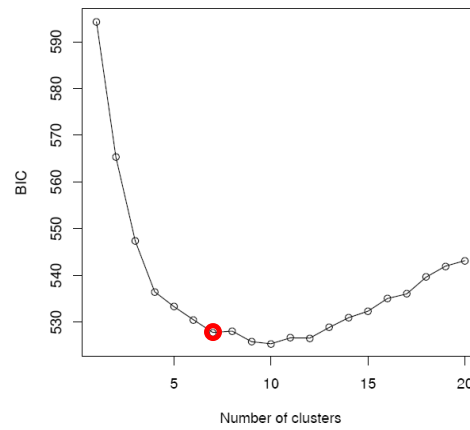


**Figure S2:** Inference of the number of clusters in the DAPC analyses performed (Discriminant analysis of principal components (DAPC, Jombart *et al.* 2010) performed **(a)** on the total global sample of 735 individuals **(b)** on South-East Asian and Oceanian samples. The Bayesian Information Criterion (BIC) is provided for different numbers of clusters and the chosen value of K is marked in red.

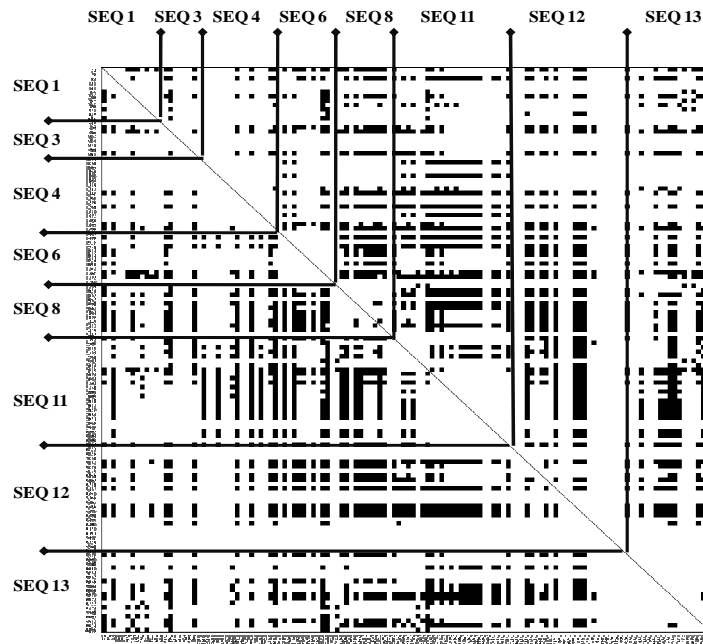
**(a)**



**(b)**



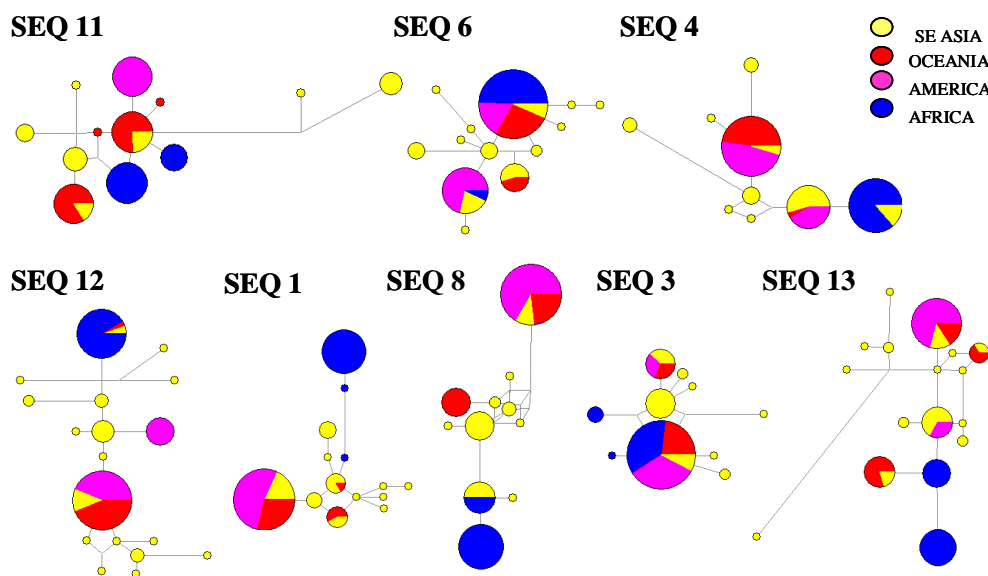
**Figure S3:** Pairwise compatibility matrix for all informative sites from all 8 sequence-based markers (excluding indels and infinite site violations), generated in SNAP Workbench on the global dataset of 120 individuals. Delimitation of the 8 markers is indicated, and solid black squares represent pairs of incompatible sites supporting conflicting phylogenies.

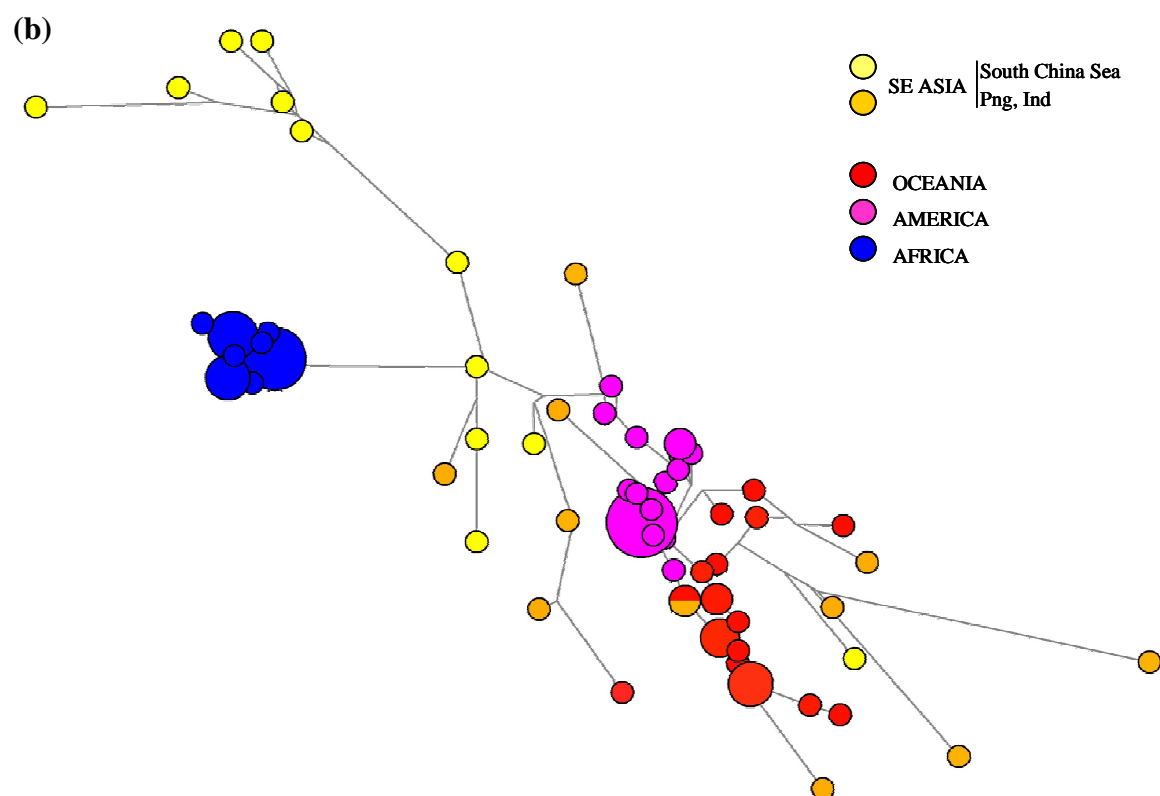


**Figure S4:** Median joining network constructed from the global sample of 120 *M. fijiensis* individuals (a) for the eight markers separately (b) for multilocus haplotypes based on combined data from all 8 sequence-based markers.

Pie size represents the number of individuals sharing the considered haplotype, and branch lengths represent the amount of genetic changes between haplotypes. Geographical origin is indicated by different colours.

(a)





**Table S1:** Reference, multiplexing information and genetic variability of the 21 microsatellite markers used in this study. Diversity indices have been evaluated on the global sample of 735 *M. fijiensis* isolates.

NA: Number of alleles; H<sub>E</sub>: gene diversity calculated following Nei (1978) in FSTAT

Multiplex pool	Locus name	Repeat motif	Reference	Labelling dye	NA	H <sub>E</sub>
1	MfSSR313	(TA) <sub>5</sub>	Zapater <i>et al.</i> 2008	NED	5	0.5
	MfSSR319	(CT) <sub>17</sub> C(CT) <sub>1</sub>	Zapater <i>et al.</i> 2008	FAM	18	0.6
	MfSSR324	(TG) <sub>14</sub>	Zapater <i>et al.</i> 2008	NED	19	0.63
	MfSSR401	(CTCA) <sub>7</sub>	Robert <i>et al.</i> 2010	PET	10	0.64
	MfSSR407	(CCAT) <sub>7</sub>	Robert <i>et al.</i> 2010	VIC	10	0.72
	MfSSR410	(CCAT) <sub>11</sub>	Robert <i>et al.</i> 2010	VIC	15	0.73
	MfSSR412	(TTGC) <sub>7</sub>	Robert <i>et al.</i> 2010	FAM	13	0.61
2	MfSSR061	(CAA) <sub>8</sub>	Neu <i>et al.</i> 1999	FAM	15	0.52
	MfSSR303	(GT) <sub>3</sub> (...)(GA) <sub>3</sub>	Zapater <i>et al.</i> 2008	NED	6	0.36
	MfSSR403	(CAAG) <sub>14</sub>	Robert <i>et al.</i> 2010	NED	13	0.76
	MfSSR406	(CTTC) <sub>6</sub>	Robert <i>et al.</i> 2010	PET	4	0.55
	MfSSR424	(TACC) <sub>5</sub>	Robert <i>et al.</i> 2010	VIC	8	0.68
	MfSSR425	(TGAT) <sub>5</sub>	Robert <i>et al.</i> 2010	PET	5	0.3
	MfSSR428	(GAAG) <sub>5</sub>	Robert <i>et al.</i> 2010	NED	3	0.47
3	MfSSR355	(GT) <sub>7</sub>	Zapater <i>et al.</i> 2008	NED	10	0.55
	MfSSR305bis	(CA) <sub>6</sub>	Zapater <i>et al.</i> 2008	NED	5	0.54
	MfSSR340	(CT) <sub>3</sub> (...)(CA) <sub>9</sub>	Zapater <i>et al.</i> 2008	FAM	5	0.4
	MfSSR362	(GA) <sub>4</sub> (...)(GT) <sub>6</sub>	Zapater <i>et al.</i> 2008	FAM	8	0.67
	MfSSR402	(TTGC) <sub>7</sub>	Robert <i>et al.</i> 2010	NED	9	0.68
	MfSSR420	(CTGA) <sub>6</sub>	Robert <i>et al.</i> 2010	VIC	6	0.33
	MfSSR430	(CTTT) <sub>5</sub>	Robert <i>et al.</i> 2010	PET	4	0.46

**Table S2:** Molecular characterization and genetic variability of the eight intron-based nuclear markers developed for this study. Diversity indices were evaluated on the global sample of 120 *M. fijiensis* isolates.

Marker name	Departure point : Gene annotation or Funybase identifier (FG)	Genome position *	Forward and reverse Primers	Outgroup amplification	Total amplified size (bp)	Annealing temperature (°C)	Final size of marker (bp)	N° of polymorphic sites	N° of haplotypes
SEQ1	CaM (Calmoduline)	2 / 1575757:1576279	TGGAGAAGCTCAGTGCACAA TCAGGTGCTATCCTCCATGA	<i>M. eumusae</i> , <i>M. musicola</i>	671	55	523	12	14
SEQ3	G3PD (Glyceraldehyde 3-phosphate dehydrogenase)	2 / 3927746:3928234	TTCATTGAGCCACACTACGC AGTAACCCCACTCGTTGTCTG	-	996	55	490	10	7
SEQ4	FG855	1 / 6056593:6056982	GGAGGTTGCACAATACAACG CGTATTCCGGCCGTCAAAG	<i>M. musicola</i>	986	55	390	16	9
SEQ6	FG612	7 / 1313554:1313944	CTCCAACAACCTCCCAAGA GCTTGGTCACCTTGTCTCTC	<i>M. eumusae</i> , <i>M. musicola</i>	563	55	391	14	12
SEQ8	FG552	2 / 1700862:1701504	CGGAAACGTGTTGTTGGAGT TCTTCGCATCCTCGTCTTCT	-	979	55.5	645	13	10
SEQ11	FG459	2 / 323952:324526	AAAGGTCCACCAGTCAATTG CTCACGAGTACGCTCACCAG	<i>M. eumusae</i> , <i>M. musicola</i>	740	55	575	25	11
SEQ12	FG1000	10 / 2344442:2344990	CCATGAAGTCAAGTCTCCA ACCCACGAGCTCACACAAAT	<i>M. musicola</i>	991	55.5	549	31	18
SEQ13	Translation elongation factor EF1 alpha	1 / 2633710:2634340	GACGGACAATGAGTCTACAA TGGCCTCGAGGAGAGTCTTA	<i>M. musicola</i> , <i>M. eumusae</i>	966	55	633	21	17

\* Scaffold/sequence position according to the *M. fijiensis* genome sequence v2.0 (<http://genome.jgi-psf.org/Mycfi1/Mycfi1.home.html>)



**Table S3:** Results of several tests run on the 25 *M. Fijiensis* populations in the software BOTTLENECK, in order to test the existence of an heterozygosity excess observed in the samples (calculated as Nei's gene diversity, Nei 1987) compared to the heterozygosity expected under mutation-drift equilibrium and termed the allelic richness. Tests were performed under two mutational models: SMM (step-wise mutation model) and IAM (Infinite alleles model).

Only p-values of significative results are represented: marginally significant ( $p < 0.1$ ), significant =  $p < 0.5$  \*, highly significant =  $p < 0.01$  \*\*, very highly significant =  $p < 0.001$  \*\*\*

Geographic area	Population	% of polymorphic loci	Sign test, IAM	Sign test, SMM	Standardized difference test, IAM	Standardized difference test, SMM	H excess Wilcoxon test, IAM	H excess Wilcoxon test, SMM
S-E ASIA	Ind	90 (19/21)	**		**	*	**	*
	MAL1	80 (17/21)						
	MAL2	71 (15/21)	*		*		*	
	PHL	95 (20/21)						
	Png1	83 (15/18)			*		*	
	Png2	100 (21/21)						
OCEANIA	FIJ	62 (13/21)				$p < 0.1$		
	Ncl	57 (12/21)						
	AUS	81 (17/21)			$p < 0.1$		$p < 0.1$	
AMERICA	COL	65 (13/20)						
	CRI	76 (16/21)	*	*	**	$p < 0.1$	**	$p < 0.1$
	H	71 (15/21)	***	***	***	***	***	***
	H-G	70 (14/20)	***	***	***	***	***	***
	MEX	71 (15/21)	***	*	***	***	***	***
	PAN	71 (15/21)	$p < 0.1$		*		*	
	VEN	55 (11/20)	$p < 0.1$		$p < 0.1$		$p < 0.1$	
	Jam	62 (13/21)						
AFRICA	RD	48 (10/21)						
	COM	29 (6/21)						
	UGA	81 (17/21)			$p < 0.1$		$p < 0.1$	
	CAM5	52 (11/21)			$p < 0.1$		$p < 0.1$	
	CAM6	57 (12/21)						
	CIV	52 (11/21)						
	Gab	57 (12/21)			$p < 0.1$		$p < 0.1$	
	NGA	38 (8/21)						

**Table S4:** Haplotype distribution of the global sample of *M. fijiensis* for the 8 sequence-based markers. Numbers in parenthesis correspond to the number of individuals sharing each haplotype.

H: haplotype; t: transitions; v: transversions, i: informative sites; -:uninformative site

	SEQ 1	SEQ 3	SEQ 4	SEQ 6	SEQ 8
Position	11222335 278240366771 493118626145	12333444 767885559 9716455781	111222222333 3446602336796778 5704731093241392	11122223333 9269001445667 30057788561064	1123345566 235956742924 2655407980585
Site Number	111 123456789012	1 1234567890	1111111 1234567890123456	11111 12345678901234	11111 1234567890123
Consensus	AGCGGCAGTTCA	CCTAGCAGTT	AGTCTCCGATCTTAAA	AGGGTCAACTGACA	CCGGTACCACTAT
Site Type	v v v v t t t v v t t i - i - i i i i i i	t t t t v v v t t t - - i - - - i -	t t t t t t t v t t t t t - - - - i i - i - - i	v t v v v t t t t t t t t - - i i - i - - i i -	v t t t v t t t t v v t t - i i - i i - - i i -
Character Type	H1 (6) . . . . . H2 (59) C . . . . . G . . . . . H3 (30) C . . . . . T . C . C . G H4 (1) . T A . . . . . A . . . . . H5 (1) . . . A . A . . . . . H6 (1) . . . A . . . . . A . . . . . H7 (7) . . . A . . . . . G . . . . . H8 (1) . . . A . . . . . A . . . . . H9 (1) . . . A . . . . . C . . . . . H10 (5) . . . C . . . . . T . . . . . H11 (1) . . . . . T . C . C . G H12 (4) . . . . . G . . . . . H13 (1) . . . . . T . . . . . H14 (1) . . . . . G . . . . .	H1 (14) . . . . . H2 (4) T . . . . . T C . . . . . H3 (12) T . . . . . H4 (81) . . C . . . . . T . . . . . H5 (1) . . . C G . . . . . H6 (1) . . . C . C A . A . . . . . H7 (2) . . . . . C . . . . .	H1 (24) . . . . . T . . . . . G . . . . . H2 (3) G . . . . . C . . . . . H3 (3) . . A C . . . . . G A . C . G . . . . . H4 (1) . . . . . T . C . . . . . H5 (1) . . . . . T . . . . . G . . . . . H6 (1) . . . . . C . . . . . H7 (43) . . . . . C . . . . . H8 (4) . . . . . H9 (37) . . . . . T . . . . . G G . . . . .	H1 (63) . . . . . G . . . . . H2 (1) T . . . . . H3 (11) A . T . . . . . H4 (1) . . T . . . . . G . . . . . H5 (2) . . . T . . . . . G . . . . . H6 (1) . . . . . G . . G . C . G . . . . . H7 (1) . . . . . G . . G . . . . . H8 (4) . . . . . T G . T . . . . . H9 (1) . . . . . G . . T . . . . . H10 (27) . . . . . T . . . . . H11 (5) . . . . . H12 (1) . . . . . T G . . . . .	H1 (47) . . . . . G . T T G . . C . . . . . H2 (26) G T A . . . G . . . . . H3 (14) . . T A . . G . . . . . H4 (1) . . T A . . G . . . . . G . . . . . H5 (11) . . . A . . . . . H6 (1) . . . A . . . . . C . . . . . H7 (1) . . . . . H8 (3) . . . . . H9 (13) . . . . . G G . . . . . H10 (1) . . . . . G . . . . . G . . . . .

	SEQ 11	SEQ 12	SEQ 13
Position	111122233333334455555 1679347734606667790246667 0731901310191395812640170	111133333333333344455555 112554499223447777888923611223 5140681547033181234837037902473	111112333356 13366899000050335593 04546628046866591892
Site Number	1111111111222222 1234567890123456789012345	111111111122222222233 1234567890123456789012345678901	11111111112 12345678901234567890
Consensus	ACACCGGTTTGCAGATCTCCATTCC	GATTATACCAACGACATCCTTAGCTAACTCG	GACTCGGACCCCTCACTGCC
Site Type	t v v t v v v v t t v v v v v t t v t t v t - i i i - i i - i - i i i i i - i i - i	v t t t t v v v t t t v v v t t v v v t t v v v - - i i - i i - i - - i i i i - - - - -	t v t t v v v t t t t t v v t t i - i - i - i i i i - i - - - i -
Character Type	H1 (7) . . . . . C C . T . . . . . T . . . . . H2 (19) G . . . . . C C . T . . . . . T . . . . . H3 (18) . . A . . . . . H4 (6) . . C . G . . . . . G . T C C C T G T . . C C . . H5 (1) . . C . G . . . C . T . . . . . T . . . . . H6 (4) . . . T . . T . C . T . . . . . T . . . . . H7 (1) . . . T . . . . . H8 (9) . . . . . A . . . . . H9 (21) . . . . . H10 (1) . . . . . C . C C T G T T T C C A . H11 (1) . . . . . H12 (21) . . . . . C . . . . .	H1 (7) . . . . . C . . . . . H2 (1) C . . . G C . . . . . C . . . C . . H3 (2) . . G . . . . . C . . . . . H4 (1) . . C . . . . . A . . . . . C T A . . . . . C . . . . . H5 (1) . . . . . G . G . . . T . C T A . . . . . H6 (1) . . . . . G . . . . . A . . . . . C . . . . . G . T . . . . . H7 (35) . . . . . G . C A . . . . . H8 (1) . . . . . G . . . . . T . . . . . C . . . . . C . . A . . . . . H9 (3) . . . . . A . . . . . G . . . . . C T A . . . . . H10 (1) . . . . . A . . . . . G . . . . . C C T A . . . . . H11 (1) . . . . . G G . . . . . C T A . . . . . G . . . . . H12 (1) . . . . . G . . . . . C T A . . . . . H13 (3) . . . . . H14 (10) . . . . . C . . . . . A . . . . . G . . . . . H15 (1) . . . . . C . . . . . G . G . . . C T A . . . . . H16 (48) . . . . . C . . . . . G . . . . . C T A . . . . . H17 (1) . . . . . C . . . . . G . . . . . H18 (1) . . . . . C . . . . . A . . . . .	H1 (15) . . . . . T T . . . . . H2 (15) A . . . . . T . . . T T . . . . . H3 (21) A . . . . . T T . . T . T . . . T . . . . . H4 (12) A . . . . . T T . . . . . H5 (2) . . T . . . . . T T . . . . . H6 (1) . . T C A . . C T . . . . . H7 (2) . . . C . . . C T . . . C . . . . . H8 (1) . . . C . A . C T . . . C . . . . . H9 (1) . . . C . . . C T . . . T C . . . C . . . H10 (1) . . . C . . . C T . . . C . C . . . . . H11 (2) . . . . . H12 (37) . . . . . H13 (7) . . . . . C . . . . . H14 (1) . . . . . T T . . . . . C . . . . . H15 (2) . . . . . T T . . . . . C T . . . . .

# OPTIMIZED GENOTYPING WITH MICROSATELLITE MARKERS IN THE FUNGAL BANANA PATHOGEN *MYCOSPHAERELLA FIJIENSIS* (MYCOSPHAERELLACEAE)<sup>1</sup>

STÉPHANIE ROBERT<sup>2</sup>, ADRIEN RIEUX<sup>2</sup>, XAVIER ARGOUT<sup>3</sup>, JEAN CARLIER<sup>2,4</sup>,  
AND MARIE-FRANÇOISE ZAPATER<sup>2</sup>

<sup>2</sup>CIRAD, Unité Mixte de Recherche–Biologie et Génétique des Interactions Plante-Parasite, TA A-54/K, Campus International de Baillarguet, F-34398 Montpellier Cedex 5, France; <sup>3</sup>CIRAD, Unité Mixte de Recherche–Développement et Amélioration des Plantes, TA A-96/03, Avenue Agropolis, F-34398 Montpellier Cedex 5, France

- *Premise of the study:* Large-scale population genetics studies are required to investigate the dispersal processes underlying the emergence of *Mycosphaerella fijiensis*, a fungal pathogen of banana. To this end, we have developed an optimized genotyping procedure combining novel microsatellite markers and a modified DNA extraction protocol.
- *Methods and Results:* Primers for tetranucleotide loci were designed directly from the recently published genome sequence of *M. fijiensis*. A total of 19 new polymorphic and easy-to-score markers were developed. Their use was combined with an adapted protocol for total DNA extraction starting from young lesions collected from banana leaves, thus avoiding a pathogen isolation step.
- *Conclusions:* The combination of the two technical developments presented here will permit the expansion of genotyping capacity in *M. fijiensis*, allowing large-scale analysis of samples from various geographic locations.

**Key words:** DNA extraction; genotyping; *Mycosphaerella fijiensis*; tetranucleotide microsatellites.

Population genetics studies on fungi are crucial to our understanding of the dispersal process of such organisms causing emerging diseases in plants. Such studies can be extended by integrating tools from landscape genetics (Archie et al., 2009). These tools require a sufficient number of neutral genetic markers and increasingly large genotyping capacities. For plant-pathogenic fungi, an important advantage is that DNA can be extracted directly from plant lesions, avoiding time-consuming isolation of the pathogen. However, this technique has seldom been used and needs to be combined with specific, easy-to-score markers to compensate for the possibly suboptimal amplification conditions.

The Ascomycete and haploid fungus *Mycosphaerella fijiensis*—the causal agent of the black leaf streak disease of bananas—can be considered as a model in the field of emerging plant diseases (Halkett et al., 2010). It has spread recently to almost all banana-producing regions worldwide and is considered among the most serious biological threats to food security. Population structure studies with microsatellite markers have been published at local scales (Halkett et al., 2010). However, most of

the microsatellite loci previously published in *M. fijiensis* (Neu et al., 1999; Zapater et al., 2008; Yang and Zhong, 2008) have been adapted for specific genetic studies at local scales, limiting comparisons among localities at continental or global scales. Indeed, most such loci correspond to di- and trinucleotides or imperfect motifs, thus increasing errors in allele identification when the number of alleles is high, as would be expected at large geographical scales. Moreover, their use is based on DNA extraction from strains isolated and cultivated on artificial medium. Consequently, to address both these issues, we have developed new microsatellite markers corresponding to tetranucleotide motifs identified in the recently published genome sequence of *M. fijiensis*. The use of these easier-to-score markers was combined with a low-cost DNA extraction protocol starting from young lesions collected on banana leaves.

## METHODS AND RESULTS

The nuclear genome sequence of *M. fijiensis* was released recently on the JGI (Joint Genome Institute, <http://genome.jgi-psf.org/Mycfi2/Mycfi2.home.html>); the sequence was analyzed for microsatellite content using the program MISA (MicroSatellite identification tool, <http://pgsc.ipk-gatersleben.de/misa>). We selected 35 tetranucleotide loci scattered throughout the genome. A total of 35 primer pairs was designed with PRIMER3 software (<http://frodo.wi.mit.edu/primer3/>, Rozen and Skaletsky, 2000), with an optimum size of 20 nucleotides, an optimum melting temperature of 60.0°C and a minimum GC content of 50%. These primer pairs were first tested on DNA from eight *M. fijiensis* isolates from Asia, Africa, and America following the procedure described in Zapater et al. (2008). About 20 loci giving rise to a single amplification product were selected on 2% agarose gels.

The polymorphism of these sequences was further tested on a global DNA sample of 110 cloned isolates from 19 countries around the world (including Taiwan, China, Vietnam, Philippines, Indonesia, Malaysia, Papua New Guinea,

<sup>1</sup> Manuscript received 6 August 2010; revision accepted 16 September 2010.

The authors thank Arnaud Estoup for helpful discussions and Helen Rothnie for her attentive reading and careful English-language checking. Data used in this work were produced through the molecular genetic analysis technical facilities of the IFR119 “Montpellier Environnement Biodiversité.” This work was supported by the région Languedoc Roussillon, Bayer CropScience, the Agence Nationale de la Recherche et de la Technologie, and the Agence Nationale de la Recherche, contract number ANR 07-BDIV-003 (Emerfundis project).

<sup>4</sup> Author for correspondence: jean.carlier@cirad.fr

Costa Rica, Honduras, Panama, Colombia, Dominican Republic, Jamaica, Cuba, Cameroon, Ivory Coast, Nigeria, Uganda, and Comoros), and two populations of 40 isolates from Cameroon and Costa Rica. All the strains were isolated and conserved in the collection of CIRAD. Total DNA was extracted from cultivated mycelium in vitro as described in Halkett et al. (2010). PCR was performed according to the following multiplex protocol (adapted from Zapater et al., 2008): 15 min at 95°C, 45 cycles of 30 s at 94°C, 90 s at 57°C and 90 s at 72°C, and a 10-min final step at 65°C. The 10-μL reaction mixture included 3 μL DNA (5 ng/μL), 5 μL (2×) QIAGEN Multiplex PCR Master Mix (Qiagen, Valencia, California, USA), 1 μL (5×) Q-solution and 0.5 μL of each primer. Forward or reverse primers were labeled with one of the following fluorescent dyes: NED, HEX, FAM, or PET (Applied Biosystems, Foster City, California, USA). Primers were tested in panels of four markers labeled with those different dyes. Diluted amplified PCR products were mixed with GeneScan-500LIZ size marker (Applied Biosystems) and separated on a 16-capillary Sequencer (ABI Prism 3130XL, Applied Biosystems). They were sized and analyzed using GeneMapper Software (Applied Biosystems); 19 loci that were technically easy to analyze were retained (Table 1).

Tetranucleotide motif markers appear particularly easy to score because alleles were, as expected, separated by 4 bp and thus were easy to identify without ambiguity across samples.

The mean number of null alleles across loci ranged from 0.9% (1/110) to 6.4% (7/110) in the global sample and from 0 to 7.5% (3/40) in the two populations. Genetic diversity was estimated by the number of alleles observed ( $N_a$ )

and the expected heterozygosity ( $H_E$ ), calculated using FSTAT<sub>2.9.3</sub> (Goudet, 1995; Table 2). At the global scale for each locus,  $N_a$  ranged between 3 and 14 (mean 7.58), and  $H_E$  from 0.31 to 0.80 (mean 0.63). Analyses were also performed within the three distinct continents: Africa and America where *M. fijiensis* was recently introduced, and Southeast Asia corresponding to its center of origin. As expected, all markers were polymorphic in Southeast Asia, with a mean  $N_a$  and  $H_E$  across loci of 6.9 and 0.74, respectively. They displayed lower levels of polymorphism in Africa and America with a mean  $N_a$  of 2.11 and 1.95, and a mean  $H_E$  of 0.30 and 0.32, respectively. Within populations from Cameroon and Costa Rica, the mean  $N_a$  and  $H_E$  varied between 1.5 to 2 and 0.18 to 0.27, respectively. No significant linkage disequilibrium among any pair of loci was detected using Fisher's exact tests in the GENEPOP program (Raymond and Rousset, 1995), and we used the false discovery rate procedure in the QVALUE package (Storey, 2002) to control for multiple testing. All markers can easily be included in multiplex panels in the coming studies. Since some markers appeared monomorphic in some areas, different sets of loci will have to be chosen for microsatellite multiplexing, depending on the geographic area studied.

Simultaneously, we adapted the DNA extraction protocol starting from in vitro grown mycelium described in Halkett et al. (2010) for young lesions directly cut out from banana leaves. Individual lesions of stage 2 (see description in Jones, 2000) were cut and dried for 1 d at 55°C. In contrast with the protocol described in Halkett et al. (2010), the extraction buffer was added in a two-step procedure allowing a better grinding of the lesions.

TABLE 1. Characteristics of the 19 tetranucleotide microsatellite loci developed for *Mycosphaerella fijiensis*.

Locus name	Primer sequences	Repeat motif	Genome position <sup>a</sup>	Expected size (bp)	Allele size range <sup>b</sup> (bp)	Annealing temperature $T_a$ (°C)
MfSSR401	F: GACGTGGCAACATCAACATC R: CTCAGCCTGTGGTGTGTTGAG	(CTCA) <sub>7</sub>	1/8304341–8304529	188	173–227	55
MfSSR402	F: CTTCTCGTCTCCACACTTCC R: CACACGAACCTTGTGCATCA	(TTGC) <sub>7</sub>	5/1559726–1559952	226	215–231	55
MfSSR403	F: TAGTGGTTGGCGGATATGGT R: AGGCTCAGCTCATCGGTAGA	(CAAG) <sub>14</sub>	10/2377024–2377220	196	162–215	55
MfSSR404	F: GCTGCTTTGGTTTGTCTAAGG R: GTCGAACCTTGCCTGAGAC	(CAGG) <sub>12</sub>	4/2435930–2436142	212	194–231	55
MfSSR405	F: TGTTCACGGAGCAAGACAAG R: GATGGTATGGCGTGGTCTCT	(CATT) <sub>6</sub>	12/824577–824693	116	101–120	55
MfSSR406	F: AGCCATTGTGTAGCGTCTTG R: GGCAATACAACCCGAGGATA	(CTTC) <sub>6</sub>	2/922431–922602	171	171–183	55
MfSSR407	F: CTCATATGGCTCAGGGTGAAT R: TGTGGAAGGAGACTTTGG	(CCAT) <sub>7</sub>	5/3498739–3498949	210	206–244	55
MfSSR410	F: GTAACGAACCTGTCTGCATA R: CCCCAATGTCGTAAACCATC	(CCAT) <sub>11</sub>	10/656074–656203	129	103–172	55
MfSSR412	F: CGGGGTCTCGGTATCAAGTA R: CGTGTTACGCTTGATTGTGG	(TTGC) <sub>7</sub>	15/603217–603389	172	145–187	55
MfSSR413	F: TGGGCCCTTACTTGTGAAC R: AACACGGGGGTACAAGACAG	(TCAC) <sub>7</sub>	7/2701378–2701597	219	201–245	55
MfSSR417	F: AGGGTGTACACATCGGCTTC R: CTTCTTGTGCAACCGCTAC	(TTCT) <sub>6</sub>	5/5244786–5244989	203	203–254	55
MfSSR418	F: GGACATGTGTGATTTGATGC R: TCGACCTGTACATCTCCAA	(GCCT) <sub>8</sub>	2/8232842–8233017	175	150–179	55
MfSSR420	F: CGAGCGAAGCTGAATAGTCC R: GTTCTGTGCTTGAGCGAGTG	(CTGA) <sub>6</sub>	7/3524840–3525030	190	182–202	55
MfSSR422	F: GAGCTGCAACCAAGAGAAGC R: GCTTCTAGCCCTGCAAACAC	(TCCA) <sub>6</sub>	6/1052203–1052412	209	196–213	55
MfSSR424	F: CGATACCCACCACATTGAT R: CTTCTGCTGAGAGCGGAATC	(TACC) <sub>5</sub>	2/4088730–4088903	173	175–194	55
MfSSR425	F: AACCACACATCTGTCTCC R: TTAGAGGTTGGATGCCGAGT	(TGAT) <sub>5</sub>	2/1374137–1374364	227	222–235	55
MfSSR428	F: AGGATTACCCGAGAACAG R: GGGTCTAGGGAATGCCTACT	(GAAG) <sub>5</sub>	4/5852974–5853108	134	127–131	55
MfSSR430	F: TAGCGAGGTGCATGGTACTG R: GGGGTTGAGGGATATTCGAC	(CTTT) <sub>5</sub>	7/4088786–4088968	182	170–186	55
MfSSR434	F: TAACGTGCGAGGAGGACTCT R: AGATCAGTGCCCCCTTGAGAA	(TACC) <sub>5</sub>	4/1802056–1802280	224	209–253	55

<sup>a</sup> Scaffold/sequence position from the forward to the reverse primers according to the *M. fijiensis* genome sequence v2.0 (<http://genome.jgi-psf.org/Mycfi2/Mycfi2.home.html>). Amplified DNA sequences can be retrieved by copying their positions to the browser of the JGI site.

<sup>b</sup> Based on a global DNA sample of 110 isolates.

F = forward primer, R = reverse primer

TABLE 2. Genetic diversity analyses at continental and intrapopulation scales in *Mycosphaerella fijiensis*, using the 19 tetranucleotide microsatellite markers.

Locus	Global scale ( <i>N</i> = 110)						Intrapopulation			
	Southeast Asia ( <i>N</i> = 32)		Africa ( <i>N</i> = 45)		Latin America ( <i>N</i> = 33)		Cameroon ( <i>N</i> = 40)		Costa Rica ( <i>N</i> = 40)	
	H <sub>E</sub>	N <sub>a</sub>	H <sub>E</sub>	N <sub>a</sub>	H <sub>E</sub>	N <sub>a</sub>	H <sub>E</sub>	N <sub>a</sub>	H <sub>E</sub>	N <sub>a</sub>
MfSSR401	0.83	8	0.24	2	0.22	2	0.49	2	0.14	2
MfSSR402	0.74	4	0	1	0.52	2	0	1	0.55	3
MfSSR403	0.85	7	0.22	5	0.48	2	0	1	0.51	2
MfSSR404	0.88	14	0.51	2	0.23	2	0.37	2	0.23	3
MfSSR405	0.85	9	0.51	2	0.52	3	0.48	2	0.63	3
MfSSR406	0.69	4	0.17	2	0.48	2	0	1	0.43	2
MfSSR407	0.84	8	0.51	2	0.47	2	0.51	2	0.50	2
MfSSR410	0.89	12	0.09	2	0.46	2	0	1	0.51	2
MfSSR412	0.86	8	0.51	2	0.31	2	0.19	2	0.41	2
MfSSR413	0.85	9	0.48	2	0.29	2	0.34	2	0.11	2
MfSSR417	0.77	6	0.51	2	0.62	4	0.51	2	0.52	4
MfSSR418	0.68	3	0.33	4	0.33	2	0.05	2	0	1
MfSSR420	0.79	6	0	1	0	1	0	1	0	1
MfSSR422	0.76	6	0	1	0	1	0	1	0	1
MfSSR424	0.83	8	0.51	2	0.52	2	0.51	2	0.10	2
MfSSR425	0.39	4	0.14	2	0.19	2	0	1	0.06	2
MfSSR428	0.41	3	0.31	2	0.49	2	0	1	0.47	2
MfSSR430	0.65	6	0.13	2	0	1	0	1	0	1
MfSSR434	0.59	5	0.50	2	0.00	1	0.05	2	0.05	2
Means	0.74	6.90	0.30	2.11	0.32	1.95	0.18	1.53	0.27	2.05

*N* = size of the sample analyzed, H<sub>E</sub> = expected gene diversity, N<sub>a</sub> = number of alleles

Dried lesions were first ground in 200 µl of 65°C extraction buffer (1.05 M NaCl, 75 mM Tris-HCl 1M pH 8.0, 15mM EDTA 0.5M pH 8.0) in 2 ml micro-centrifuge tubes, then 100 µl of 65°C cetyltrimethylammonium bromide (CTAB) buffer (3% CTAB, 1.5% sodium sulfite) was added and the tubes were placed at 65°C for 60 min. The other steps of the extraction protocol described in Halkett et al. (2010) remained unchanged.

A selection of 11 new microsatellite markers polymorphic at the scale of Cameroon were successfully amplified and analyzed from a sample of 60 DNA samples extracted from lesions originating from Cameroon. For each locus, the same allele sizes and a similar number of null alleles (range from 0 to 5% (3/60) across loci) were detected in comparison with the Cameroonian sample of cloned isolates. However, multiple allele patterns (generally with two alleles) for several loci were detected for 13.3% (8/60) of the DNA samples. Since *M. fijiensis* is haploid, these patterns certainly came from the co-alescence of lesions initiated by different spores. Such samples must be considered as missing data since multiloci genotypes could not be determined. However, this drawback could be minimized through selection of lesions sufficiently isolated from each other on leaves, and the rapid, low-cost extraction method presented here will easily allow increased sample sizes to compensate for missing data.

## CONCLUSIONS

The combination of the two technical developments presented in this note allows us to optimize our genotyping capacities in *M. fijiensis* for analysis of large samples. The tetranucleotide microsatellites developed are technically easy to score from local to global geographic scales, and the possibility of using a common set of microsatellites at different locations opens up new prospects of analysis to understand the dispersal processes in *M. fijiensis*. Similar procedures could be employed in other fungal pathogens of plants.

## LITERATURE CITED

- ARCHIE, E. A., G. LUIKART, AND V. O. EZENWA. 2009. Infecting epidemiology with genetics: A new frontier in disease ecology. *Trends in Ecology & Evolution* 24: 21–30.
- GOUDET, J. 1995. FSTAT (Version 1.2): A computer program to calculate F-statistics. *The Journal of Heredity* 86: 485–486.
- HALKETT, F., D. COSTE, G. G. RIVAS PLATERO, M. F. ZAPATER, C. ABADIE, AND J. CARLIER. 2010. Genetic discontinuities and disequilibria in recently established populations of the plant pathogenic fungus *Mycosphaerella fijiensis*. *Molecular Ecology* 19: 3909–3923.
- JONES, D. R. 2000. Diseases of banana, abaca and enset. CABI Publishing, Oxon.
- NEU, C., D. KAEMMER, G. KAHL, D. FISHER, AND K. WEISING. 1999. Polymorphic microsatellite markers for the banana pathogen *Mycosphaerella fijiensis*. *Molecular Ecology* 8: 523–525.
- RAYMOND, M., AND F. ROUSSET. 1995. An exact test for population differentiation. *Evolution; International Journal of Organic Evolution* 49: 1280–1283.
- ROZEN, S., AND H. J. SKALETSKY. 2000. Primer3 on the WWW for general users and for biologist programmers. In S. Krawetz and S. Misener [eds.], *Bioinformatics methods and protocols: Methods in molecular biology*, 365–386. Humana Press, Totowa, New Jersey.
- STOREY, J. D. 2002. A direct approach to false discovery rates. *Journal of the Royal Statistical Society. Series B, Statistical Methodology* 64: 479–498.
- YANG, B. J., AND S. B. ZHONG. 2008. Fourteen polymorphic microsatellite markers for the fungal banana pathogen *Mycosphaerella fijiensis*. *Molecular Ecology Resources* 8: 910–912.
- ZAPATER, M. F., M. DUCHEMIN, J. F. DUSSART, D. COSTE, P. BROTTIER, AND J. CARLIER. 2008. Microsatellite markers for the fungal banana pathogens *Mycosphaerella fijiensis*, *Mycosphaerella musicola* and *Mycosphaerella eumusae*. *Molecular Ecology Resources* 8: 1121–1125.



## CHAPITRE 2





## CHAPITRE 2

### **Histoire d'invasion en Afrique et en Amérique: tests des scénarios d'introduction et caractérisation de la propagation intracontinentale**

Le Chapitre 1 suggère qu'un évènement de fondation unique serait à l'origine de l'invasion du continent africain, tandis que le continent américain aurait été envahi par des populations résultant de mélange entre différentes sources, introduites probablement initialement au Honduras. Ces scénarios sont déjà suggérés de façon robuste par l'emploi combiné des microsatellites et des séquences, mais elles ne permettent cependant pas de les tester dans un cadre statistique, en les confrontant notamment à d'autres hypothèses. De plus les méthodes précédentes présentent certaines limites qui peuvent s'avérer problématiques spécifiquement pour l'étude des invasions. En particulier elles ne prennent pas explicitement en compte l'histoire démographique des populations qui peut être particulièrement stochastique, et fortement impacter les patrons de diversité génétique observés après établissement des populations (Estoup & Guillemaud 2010). Notons de plus qu'elles ne permettent pas de faire la différence entre des patrons de différenciation génétique dus à la distance géographique et aux évènements de dispersion, et ceux dus à l'isolement temporel, ce qui pourrait pourtant être particulièrement utile dans le cas d'un échantillonnage espacé dans le temps comme le notre.

Ce deuxième chapitre propose de tester formellement les scénarios d'introduction suggérés en Afrique d'une part et en Amérique d'autre part, et de préciser les mécanismes de dissémination à l'échelle continentale, en se basant entre autres sur l'emploi des approches plus récentes d'ABC (« Approximate Bayesian Computation ») qui ont l'avantage de prendre en compte explicitement les évènements démographiques qui ont pu affecter la diversité génétique des populations envahissantes. Elles nous permettront également d'estimer des paramètres démographiques associés aux scénarios testés, notamment des tailles efficaces de populations et la sévérité des goulots d'étranglement subis par les populations envahissantes au cours de la dissémination.

Ce chapitre repose sur deux grandes études focalisées respectivement sur le cas de l'Afrique et celui de l'Amérique. L'étude africaine est présentée sous la forme d'un article non publié. Concernant l'Amérique, nous présentons seulement les analyses ABC que nous avons réalisées, ainsi qu'un complément concernant les analyses préliminaires menées en équipe au niveau des Caraïbes, zone de progression la plus récente de la maladie. A terme une publication devra intégrer ces deux aspects, mais des analyses plus approfondies, notamment au niveau des Caraïbes, devront d'abord être menées.



# **Single introduction followed by multiple drastic bottlenecks characterize the emergence and spread of a banana wind-born pathogen across Africa**

S. ROBERT\*, M-F. ZAPATER\*, J. CARLIER\* et V. RAVIGNE\*

\* CIRAD, UMR BGPI, Campus international de Baillarguet, TA A-54K, F-34398 Montpellier Cedex 5, France

## **ABSTRACT**

Reconstructing geographic and demographic scenarios of dissemination in wind-dispersed invasive pathogens of crops is still poorly investigated. Here, we aimed at deciphering history of invasion of the fungal banana pathogen *Mycosphaerella fijiensis* in Africa, including initial introduction into the continent, subsequent pathways of continental spread and demographical events associated. We analysed the Africa population structure based on microsatellites markers, on 9 populations from 7 eastern and western African countries. We carried out ABC analyses in order to test different competing scenarios of between-population historical divergences, and to estimate demographical parameters describing effective size variations during the African invasion. Our results strongly support a scenario of single introduction event into Africa, and they also suggest that it was followed by both regular and stochastic events of jumps of the disease across the continent. Huge effective population sizes were estimated for established populations, which contrasts with severe past bottlenecks, reflecting that drastic demographical fluctuations characterize this continental stochastic spread.

## **INTRODUCTION**

To understand how introductions of species in new areas may successfully result in biological invasions, a precise knowledge of invasion routes and past demographic events undergone by invading populations is required (Facon *et al.* 2006, Estoup & Guillemaud 2010). Unraveling scenarios of introductions and detecting stochastic events such as bottlenecks and multiple introductions, may however be challenging. This is particularly true when considering emerging diseases of plants resulting from introductions of pathogens in new areas, for which various dissemination modes can occur even on long distances (Anderson *et al.* 2004). As



many crop and tree diseases are caused by wind-dispersed fungi, the relative contributions of natural vs. anthropogenic dissemination often remain unclear (Desprez-Loustau *et al.* 2007).

Various scenarios of introductions have been uncovered in fungal plant pathogens, from single introduction and invasions of large areas by single lineages (Engelbrecht *et al.* 2004, Raboin *et al.* 2007), to multiple introductions (Gladieux *et al.* 2008, Dutech *et al.* 2010) likely to lead to admixture between invasive populations (Zhou *et al.* 2007, Montarry *et al.* 2010). Following introduction, intra-continental spread has also been studied using analyses of population genetic structure. Most frequently, less genetic differentiation is found within continents than between continents, resulting from historical colonization processes or sometimes from occurrence of more gene flow at lower geographical scale. The contribution of wind-dispersed spores has generally been suspected, as in the case of poplar rust fungus *Melampsora larici-populina* (Barres *et al.* 2008) where an isolation-by-distance pattern was highlighted, reflecting a gradual front of dissemination. Human transports have also been incriminated, notably in wheat pathogens *Phaeosphaeria nodorum* and *Mycosphaerella graminicola* (Banke & McDonald 2005, Stukenbrock *et al.* 2006) where little structure is found within continents, probably favored by widespread exchanges of infected seeds.

Overall, the occurrence of bottlenecks at different geographical scales is a common feature of most described examples of fungal pathogens of plants such as *Rhynchosporium secalis* on barley (Zaffarano *et al.* 2009), *Mycosphaerella graminicola* on wheat (Stukenbrock *et al.* 2007), *Venturia inaequalis* on apple (Gladieux *et al.* 2008) or *Melampsora larici-populina* (Barres *et al.* 2008), but precise knowledge regarding demographic features of populations in invasive fungal organisms is still lacking. Processes of range expansion are expected to be key determinants of demographical fluctuations undergone by invading populations, as invasions are mainly characterized by complex demographic scenarios with frequent successions of bottlenecks, population divergence and expansion (Estoup & Guillemaud 2010). Nevertheless, to our knowledge, no study has tried to estimate demographic parameters describing these widespread events in fungal pathogens of cultivated plants. Many different factors can in fact influence recent demographical variations of fungal pathogens populations in cultivated plants, such as the biology of host plant (annual or perennial) and the structure of agro-systems. The few studies having focused on demographical history of crop pathogens have for example shown population expansions associated with the advent of new agricultural practices as for example in the barley scald pathogen *Rhynchosporium secalis* (Zaffarano *et al.*



2008) or the fungal wheat pathogen *Mycosphaerella graminicola* (Stukenbrock *et al.* 2007). Large uniform crops are thought to have favored high population sizes in cereals fungal pathogens. Zhan & McDonald (2004) have notably estimated a population effective size of 24000 in *Mycosphaerella graminicola*, but this remains a rare example of quantitative estimation of effective population sizes in fungal plant pathogens.

The recent worldwide emergence of the fungal pathogen of banana *Mycosphaerella fijiensis* responsible for the Black Leaf Streak Disease (BLSD) is a suitable model to address such issues. BLSD is the most destructive leaf disease of banana and a few historical data have been recorded along its spread throughout the tropical world for the last 50 years (Jones 2000). Population genetics studies revealed that, despite its wind-dispersed ascospores and its widespread sexual reproduction, *M. fijiensis* is among the most genetically structured fungal pathogen of crops, even at continental or country scales (Carlier *et al.* 1996, Rivas *et al.* 2004, Robert *et al.* 2012).

In Africa, an original scenario of introduction was uncovered thanks to the combination of clustering analyses on microsatellite markers and phylogenetic networks approaches on sequence markers. Introduction of a reduced number of genotypes coming from a single South East Asia source population was strongly suggested to have led to the invasion of the entire continent (Robert *et al.* 2012). This result allowed questioning the different hypotheses derived from historical observations, which suggested either a single or multiple introductions at different possible places, but strongly arguing in favour of a Gabonese introduction (Frossard 1980, Mourichon & Fullerton 1990). A formal test of the hypothesis of single introduction in a probabilistic framework is still lacking, as well as knowledge of its place. Regarding subsequent intra-continental dissemination, historical data suggested that the disease would have radiated from Gabon to the whole western Africa following a gradual geographical propagation front (Mourichon & Fullerton 1990). By contrast previous population genetics studies rather suggested that African range expansion had been stochastic, with frequent bottlenecks accompanying populations establishment (Rivas *et al.* 2004, Robert *et al.* 2012). Finally little is known about the intensity of these demographic events and demographic features of *M. fijiensis* populations in general. Huge census population sizes can be observed in the field with several thousand lesions per banana leaf (Gauhl *et al.* 1994), due to both sexual (heterothallic) and asexual multiplication, and the perennial banana host is thought to allow a relative demographic stability in established populations. However we have





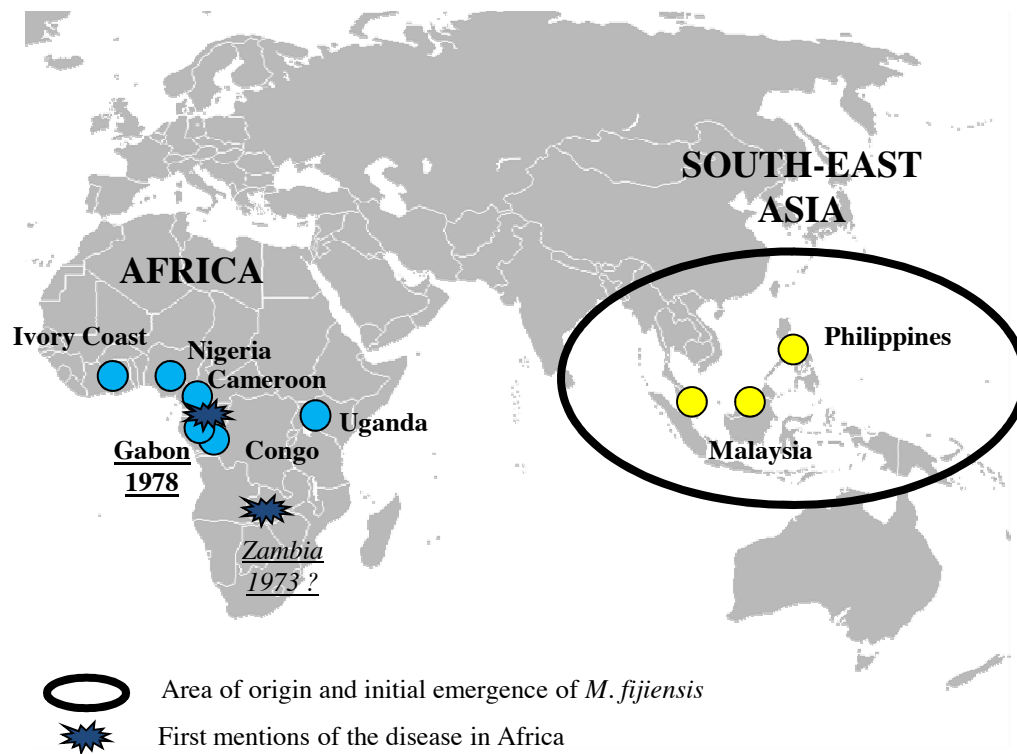
no quantitative estimations of the intensity of fluctuations of population effective sizes associated with this recent history of invasion.

In the recent years, new methodologies, based on Bayesian statistics, have arisen that allow testing hypotheses about population demographic history and estimating parameters (Bertorelle *et al.* 2010, Guillemaud *et al.* 2010). Here we aimed at using these to decipher both the pathways of African introduction and spread, and the demographical history of invading populations. Specifically, we first aimed at formally testing the scenario of two independent eastern and western African introductions vs. single introduction. Secondly, we wanted to test whether the original introduction could have taken place in Gabon in line with historical data. Thirdly, we tried to unravel propagation pathway across the continent and see whether the disease could have spread from Gabon to the whole West Africa following regular geographical propagation. Last, we looked for estimates the intensity of effective sizes of *M. fijiensis* populations and their historical variations associated with African introduction and subsequent propagation?

Eight African populations from 6 different counties have been genotyped with 21 microsatellite markers and analyzed for genetic diversity and structure with descriptive statistics and Bayesian clustering methods. Based on these results, the different concurrent scenarios for addressing our whole questions were designed and tested through two successive rounds of Approximate Bayesian Computation (ABC) analyses. Effective population sizes and intensity of bottlenecks was also assessed thanks to estimations of demographical parameters. Our results show how introductions and dissemination scenarios deciphered by genetic analyses can reveal different from historical hypotheses. Firstly we formally confirmed that a single successful introduction has invaded both the East and West Africa. But against the strongest historical hypothesis, we rejected the occurrence of this initial founder event in Gabon, and suggested a more central or eastern introduction, followed by both regular and unordered stochastic dissemination across the continent. This was accompanied by severe estimated bottlenecks at both the entrance into the continent and the subsequent spread.

## **MATERIALS AND METHODS**

### **Sampling and microsatellite genotyping**



**Figure 1: Sampling location of *M. fijiensis* populations used in this study, and historical hypotheses of introduction into Africa.**

*M. fijiensis* samples used in this study (Figure 1) were derived from a laboratory collection preserved in glycerol at -80°C. The 11 *M. fijiensis* populations were sampled at different dates and originated from Africa (8 populations from Uganda, Congo, Gabon, Cameroon, Nigeria and Ivory Coast) and South-East Asia around South China Sea (3 populations from Philippines and Malaysia). All populations represented individuals isolated simultaneously from different plants of the same banana field. Banana hosts were susceptible clones belonging to the AAA and AAB genome groups.

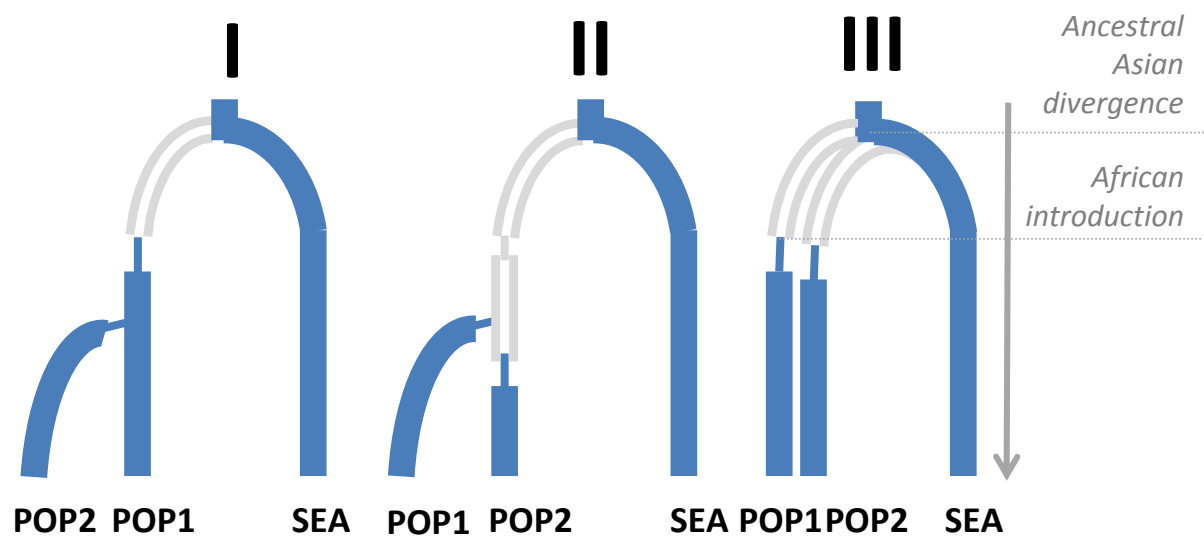
Each individual corresponds to a fungal strain derived from a single ascospore isolated from infectious banana leaf tissue. Strains rearing, DNA extraction and microsatellite genotyping with 21 microsatellite markers were realized as described in Robert *et al.* (2012).

### **Descriptive population genetics analyses**

Gametic linkage equilibrium between all pairs of loci was tested in all populations using Fisher's exact tests implemented in GENEPOP<sub>4.0</sub> (Raymond & Rousset 1995; Rousset 2008). We used the false discovery rate (FDR) procedure implemented in the R package QVALUE (Storey 2002) to control for multiple testing. Gene diversity  $H_E$  was estimated following Nei (1978), using FSTAT<sub>2.9.3</sub> (Goudet 1995). To analyze allelic diversity we used the rarefaction method implemented in HP-Rare (Kalinowski 2005) to compute allelic richness  $A_r$  and private allelic richness  $pA_r$  while correcting for samples sizes differences. Pairwise  $F_{st}$  (Weir & Cockerham 1984) between African populations were computed in FSTAT.

### **Bayesian clustering analyses**

Spatial genetic structure among African samples was examined with two different methods of Bayesian clustering. We first used the model-based procedure implemented in STRUCTURE (Pritchard *et al.* 2000), to assign individual genotypes to a pre-determined number of clusters  $K$ , minimizing linkage disequilibrium between loci within clusters. We conducted a series of 10 independent runs for  $K$  varying from 1 to 10, under a model with admixture. Each run used  $10^6$  iterations after a burn-in of  $3 \cdot 10^4$ . MCMC convergence and consistency among the 10 replicated runs were examined for each value of  $K$ . The number of clusters, above which likelihood is no longer improved, was determined following the approach of Pritchard *et al.* (2000).



**Figure 2: Scenarios under comparison in the 3-scenarios ABC analysis.**

SEA represents a South East Asian sampled population ; POP1 and POP2 represent two African sampled populations. Thin lines reflect effective population size reductions (bottlenecks). Grey lines represent unsampled populations.

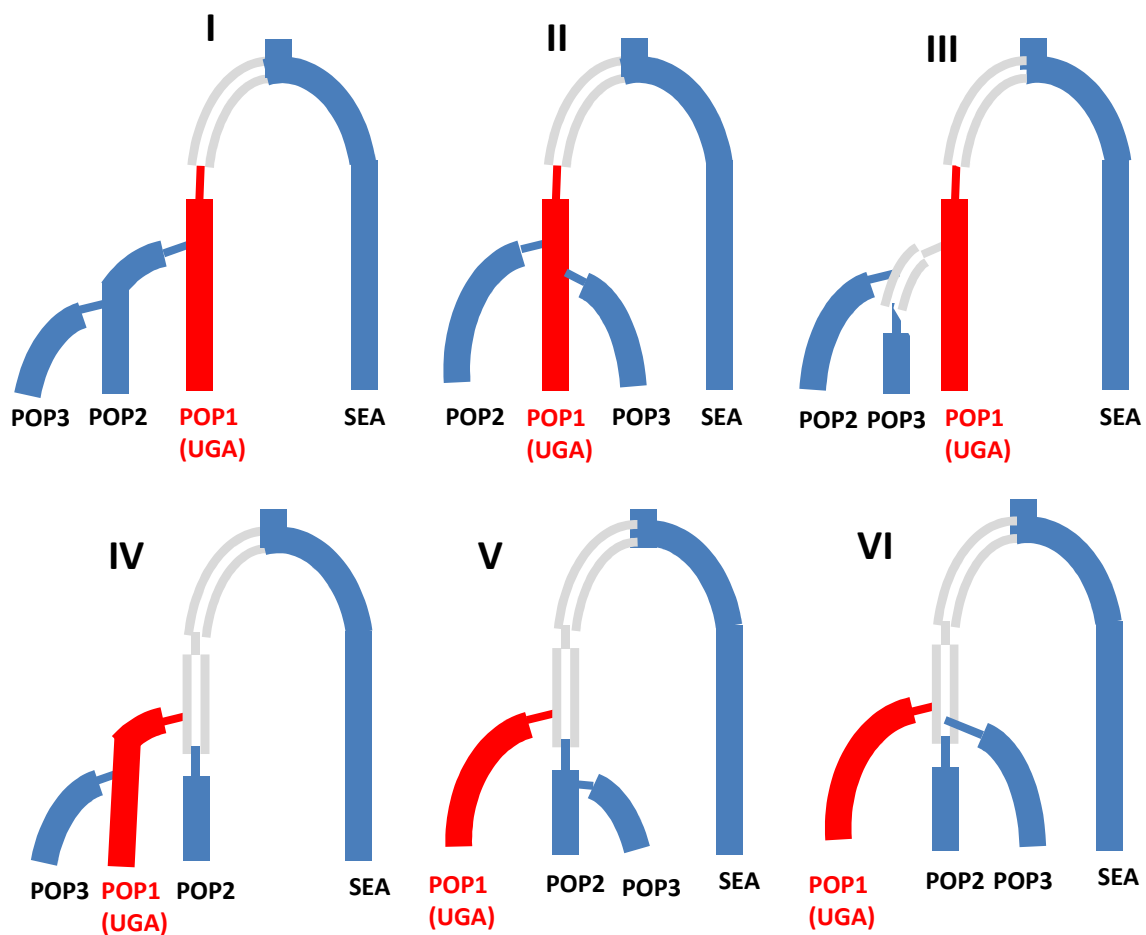
To complement this approach, we performed a discriminant analysis of principal components (DAPC, Jombart *et al.* 2010), which is model-free and allows determining the most likely number of clusters. It was performed in the R package Adegenet (Jombart 2008). Under this method the number of clusters that best describes the data is inferred using sequential K-means and the BIC criterion. Inferred clusters are then described using PCA (Principal Component Analysis) as a prior step to DA (Discriminant Analysis) allowing maximizing the separation between clusters.

### **ABC analyses with DIYABC**

ABC analyses were carried out in order to compare different scenarios of population divergence in Africa, based on their posterior probabilities, and to estimate demographic parameters associated with the best scenario. Every step of ABC method was performed with the program DIYABC (Cornuet *et al.* 2008, 2010).

#### **Testing introduction into Africa: scenarios under comparison**

In a first round of analyses, our aim was to statistically test the hypotheses of single vs. several independent introductions into Africa, and if possible, precise the location of the putative introductions. To do this, we compared 3 scenarios of population divergence including a South-East Asian population (SEA) and two African populations (POP1 and POP2, Figure 2). Given the unclear location of the real source population in South-East Asia (Robert *et al.* 2012), we considered that African populations (POP1 and POP2) were not directly founded from the sampled SEA population, but from another, still unsampled, SEA population (in grey, Figure 2). Scenario I corresponds to a single introduction in one of the populations sampled in this study (i.e., introduction occurs in POP1 through a bottleneck and POP2 derives from POP1 through an additional founding event). Scenario II models a single introduction in a population not available as a sample in the present study. Scenario III corresponds to the historical hypothesis of two separate introductions in the two sides of the continent (i.e., two different introductions of POP1 and POP2 derive from two independent SEA unsampled populations though bottlenecks). Different sampled African populations were alternatively used as POP1 and POP2, based notably on geographical considerations. DIYABC assumes an absence of migration between populations. Following clustering results, the group of populations GAB1, GAB2, CAM5, CAM6 and NGA (hereafter referred as WEST-group) were too close both geographically and genetically to exclude the possibility of



**Figure 3: Scenarios under comparison in the 6-scenarios ABC analysis.**

SEA represents South East Asian sampled population ; POP1 corresponds to UGA in all analyses. POP2 and POP3 represent African sampled populations. Thin lines reflect effective population size reductions (bottlenecks). Grey lines represent unsampled populations.

regular migration between these populations. So, in the respect of DIYABC assumptions, they were all used alternatively but never included together in the same scenario. On the contrary, CIV, COG and UGA revealed genetically distinct enough, and geographically or temporally isolated enough from the rest of populations, to consider them independently from WEST-group. Scenarios describing the divergence UGA/WEST-group, CIV/WEST-group, COG/WEST-group, UGA/CIV, UGA/COG and CIV/COG have thus been considered in our analyses. The three sampled SEA populations from Philippines and Malaysia were also used alternatively in all series of analyses in order to check for consistency of results.

#### Testing propagation pathways within Africa: scenarios under comparison

In a second round of analyses, we explored historical relationships between African populations after initial introduction, by including 3 African populations POP1, POP2 and POP3. The design of the competing scenarios was based on the results of 3-scenarios analyses (see results section) in order to exclude improbable scenarios of divergence and thus reduce the number of scenarios considered to a tractable analysis. Scenarios of divergence with the worst posterior probabilities were excluded to obtain a total of 6 competing scenarios summarized in Figure 2b. Following 3-scenarios analyses, only the two possibilities of introduction - via UGA (scenarios I to III, Figure 3) and via an unsampled population preceding UGA (scenarios IV to VI) - were considered to describe initial African introduction, with the systematic inclusion of UGA as POP1 (in red, Figure 3). Regarding the subsequent divergence between POP2 and POP3, scenario I represents the hypothesis of a regular geographical propagation, where populations POP1, POP2 and POP3 are founded consecutively. More stochastic population foundations are represented by scenarios II to VI, corresponding to the design of all possibilities of POP2 and POP3 divergence. Different African populations were also used alternatively as POP2 and POP3 (i.e., WEST-group populations, COG and CIV).

#### Prior distributions of parameters

Given our limited knowledge about demographical features of *M. fijiensis* populations, uniform distributions with large intervals [100-100000] were chosen for all effective population sizes ( $N$ ). Within Africa, the potential occurrence of bottlenecks at each population divergence was taken into account by modelling effective size variations. Uniform priors distributions over the interval [1-500] were used for effective size during bottleneck ( $N_b$ ).





Uniform distribution over the interval [0-50] and [0-200] were used for duration of bottleneck ( $d_b$ ) in the 3-scenarios and 6-scenarios rounds of analyses respectively.

Successive times of founding of sampled populations were expressed in numbers of *M. fijiensis* generations backward in the time. As ascospores are produced after around 30 to 40 days in favourable conditions (Churchill 2010), we made the approximation that *M. fijiensis* displays an average of 10 successive generations per year. Based on historical marks, we used uniform distributions over the intervals [300-100000] for ancestral divergence between SEA sampled and unsampled populations, [300-1000] for initial African introduction(s), and [150-1000] for subsequent population divergences within Africa. Finally parameters for microsatellite mutation models followed a generalized stepwise mutation model, with a log uniform distribution over the interval [ $2.10^{-6}$  -  $1.10^{-4}$ ].

#### Posterior probabilities of scenarios

One million datasets were simulated under each scenario. ABC methods rely on the comparison of observed and simulated datasets on the basis of summary statistics. In this study we have chosen to use the mean number of alleles, mean genic diversity, mean size variance and mean Garza-Williamson's M index (Garza & Williamson 2001), as well as  $F_{ST}$  (Weir & Cockerham 1984), and genetic distances ( $(\delta\mu)^2$ ) (Goldstein *et al.* 1995), taken between pairs of populations. The posterior probabilities of the four concurrent scenarios were estimated by performing a polychotomous weighted logistic regression on the 1% of simulated data sets closest to the observed data set, following a factorial discriminant analysis (FDA) on the chosen summary statistics. The selected scenario is that with the highest posterior probability value with a non-overlapping 95% confidence interval.

#### Parameter estimation

Posterior distribution of simple and composite parameters was estimated under the selected scenario, using a local linear regression on the 1% closest simulated datasets with a logit transformation (Beaumont *et al.* 2002).

#### Assessment of the validity of ABC analyses

The validity of ABC analyses was assessed using three complementary tools: computation of type I and type II error rates, model checking and precision of parameter estimates.

Firstly, to assess the confidence of our selected scenario, type I and type II errors on scenario choice were computed with DIYABC for all analyses, following the method described in

**Table 1: Genetic diversity of *M. Fijiensis* populations estimated on microsatellites. Populations were sampled in different countries and at different dates.**

Geographic area	Country	Population	Sampling date	Sample size	$H_E$	Average $H_E$	$Ar$	Average $Ar$	$pAr$
SE ASIA	Malaysia	<b>MAL1</b>	1995	13	0.41		2.7		0.34
	Malaysia	<b>MAL2</b>	1995	22	0.31	0.43	2.1	3.1	0.15
	Philippines	<b>PHL</b>	1993	22	0.58		4.4		0.61
AFRICA	Uganda	<b>UGA</b>	1998	34	0.27		1.9		0.07
	Cameroon	<b>CAM5</b>	2001	21	0.19		1.6		0.05
		<b>CAM6</b>	2001	24	0.16		1.6		0.01
	Ivory Coast	<b>CIV</b>	1999	34	0.13	0.16	1.5	1.6	0
	Gabon	<b>GAB1</b>	1998	10	0.06		1.1		0
		<b>GAB2</b>	1998	10	0.16		1.5		0
	Congo	<b>COG</b>	2011	32	0.18		2.5		0.15
	Nigeria	<b>NGA</b>	1999	24	0.10		1.4		0

$H_E$ : gene diversity;  $Ar$ : allelic richness;  $pAr$ : private allelic richness

**Table 2: Pairwise  $F_{st}$  (Weir & Cockerham 1984) estimated for all pairs of African *M. Fijiensis* populations.**

	UGA	CAM5	CAM6	CIV	COG	GAB1	GAB2
<b>CAM5</b>	0.10						
<b>CAM6</b>	0.11	0.17					
<b>CIV</b>	0.18	0.16	0.21				
<b>COG</b>	0.24	0.33	0.32	0.35			
<b>GAB1</b>	0.14	0.36	0.33	0.42	0.38		
<b>GAB2</b>	0.17	0.01	0.29	0.22	0.39	0.50	
<b>NGA</b>	0.12	0.23	0.34	0.30	0.38	0.30	0.33

Cornuet *et al.* (2010). One hundred test datasets were simulated under each of the competing scenarios (3 scenarios in the first-round and 6 in the second-round analyses), by drawing parameter values into prior distributions. For each test dataset, posterior probabilities of the competing scenarios were computed as described above. Type I error was computed as the proportion test datasets simulated under the best scenario for which the best scenario had not the highest posterior probability. Type II error was computed as the proportion of tests datasets simulated under other scenarios and for which the best scenario had the highest posterior probability.

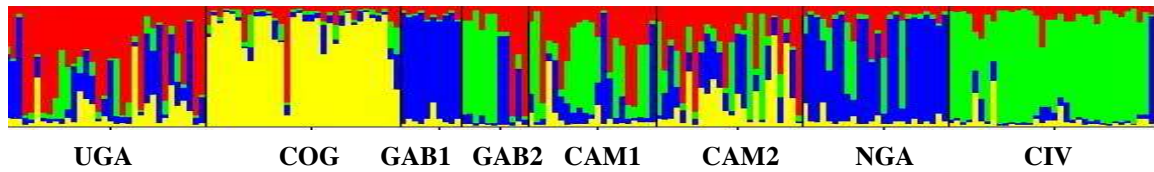
Secondly, the adequacy of the best scenario with the observed data was tested using the model-checking option of DIYABC. The principle is to assess the similarity of the real dataset with 1000 test datasets simulated under the best scenario and posterior distributions of parameters. Model checking was performed using all summary statistics provided by DIYABC that were not already used for scenario choice (mean number of alleles, mean genic diversity, mean size variance, classification index and shared allele distance (DAS, Chakraborty & Jin 1993), all taken for pairs of populations). This precaution was taken to avoid over-estimating the quality of the fit by using the same statistics twice.

Finally precision of parameter estimation was evaluated by computing the Relative Median of the Absolute Error (RMAE, Cornuet *et al.* 2010) on 500 pseudo-observed datasets simulated under the best scenario and with known parameters.

## RESULTS

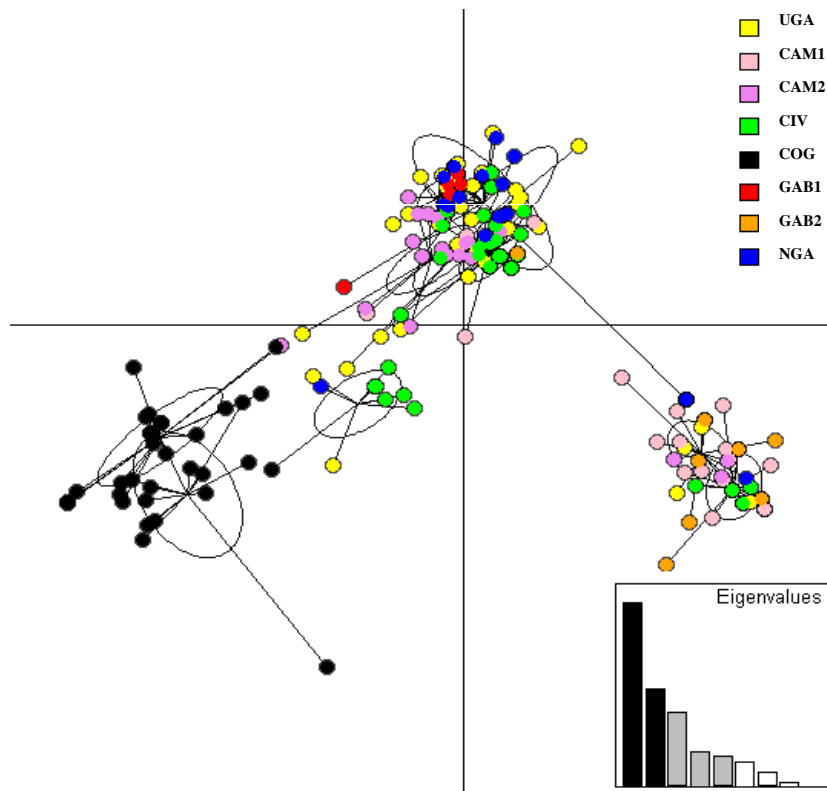
### Descriptive analyses

No significant linkage disequilibrium between any pairs of loci was noted after FDR control. Average genetic diversity was drastically reduced in Africa (average  $H_E$  : 0.16, average  $Ar$  : 1.6) as compared to South-East Asia (average  $H_E$  : 0.43, average  $Ar$  : 3.1), with very few private alleles, as previously shown (Robert *et al.* 2012). UGA was the most diverse African population ( $H_E$ : 0.27;  $Ar$ : 1.9;  $pAr$ : 0.07). The highest private richness was in COG population ( $pAr$ : 0.15) but new alleles could actually have appeared during the hundred generations separating between all other populations sampling and COG sampling (Table 1). Based on pairwise  $F_{st}$ , the African populations studied were highly differentiated (Table 2, except for the pair GAB2/CAM5), whatever their geographical distance, due to modifications of allelic



**Figure 4: Bayesian clustering of African multilocus microsatellite haplotypes, using STRUCTURE (K=4) .**

Black lines separate individuals from the different populations of origin, as indicated by population abbreviations under the barplot (detailed in Table 1).



**Figure 5: Scatterplot of Discriminant Analysis of Principal Components (DAPC, Jombart et al. 2010) performed on African multilocus microsatellite haplotypes for 10 clusters.**

Individuals are represented by dots and clusters as inertia ellipses. Colours reflect the different populations of origin indicated in the legend.

frequencies. No association between geographical distances and genetic properties of populations could be observed in these African samples.

### **Clustering analyses**

In STRUCTURE analysis (Figure 4), the most appropriate number of clusters according to Pritchard *et al.* (2000) method was  $K = 4$ . A mix between the four clusters was observed within all countries (which remains true whatever the value of  $K$ ). However, COG and CIV distinguished clearly from the other populations because for each of these, populations all individuals were preferentially assigned to one cluster.

A similar pattern was revealed by DAPC (Figure 5). The most likely number of clusters was 10. Most clusters were not well discriminated from each other on the main discriminant axes, and for most populations individuals were assigned to several clusters. Only COG was clearly separated from other African populations according the first discriminant axes.

### **ABC analyses with DIYABC**

#### **Introduction into Africa**

*Posterior probabilities.* Posterior probabilities of scenarios for the 3-scenarios analysis are reported in Table 3. Scenarios with two independent African introductions systematically received a null or negligible support, whatever the populations used. The best-supported scenarios always corresponded to two different possibilities only: scenarios I with introduction in UGA, or most frequently, scenario II modelling introduction via an unsampled African population. When UGA was included in analyses, introduction via UGA displayed the highest posterior probability in 4/7 analyses (UGA/GAB1, UGA/CAM5, UGA/CAM6 and UGA/NGA). In addition, UGA stemming from another sampled African population was always excluded. Finally, an introduction in WEST-group, CIV or COG was never the best scenarios and generally obtained low or moderate statistical support. These results altogether suggested that a single introduction has occurred in an unsampled locality, which then independently gave rise to UGA on the one hand and the WEST-group one the other hand. CIV and COG position remained uncertain (always deriving from an unsampled population, with nevertheless substantial support for a divergence of CIV from UGA).

**Table 3: Posterior probabilities of scenarios in the 3-scenarios analysis, for different alternative analyses using different populations as POP1 and POP2.**

The confident interval for each scenario is given into brackets. Scenarios I to III are exposed in Figure 2. Results are shown for SEA corresponding to PHL (Philippines population). WEST-group populations are indicated in blue. For each analysis, the selected scenario is highlighted in bold.

POP SEA	POP1 Africa	POP2 Africa	Competing scenarios (*selected scenario highlighted)	Probability of scenario (logistic regression)	Confident intervals
PHL	GAB1	UGA	I (SEA -> GAB1 -> UGA)	0.0614	[0.0575,0.0652]
			I (SEA -> UGA -> GAB1)	<b>0.6175</b>	<b>[0.6083,0.6268]</b>
			II	0.3211	[0.3124,0.3298]
			III	0.0000	[0.0000,0.0000]
PHL	GAB2	UGA	I (SEA -> GAB2 -> UGA)	0.1271	[0.1211,0.1330]
			I (SEA -> UGA -> GAB2)	0.3033	[0.2951,0.3116]
			II	<b>0.5696</b>	<b>[0.5608,0.5783]</b>
			III	0.0000	[0.0000,0.0000]
PHL	CAM5	UGA	I (SEA -> CAM5 -> UGA)	0.1283	[0.1218,0.1348]
			I (SEA -> UGA -> CAM5)	<b>0.4979</b>	<b>[0.4868,0.5089]</b>
			II	0.3738	[0.3636,0.3841]
			III	0.0000	[0.0000,0.0000]
PHL	CAM6	UGA	I (SEA -> CAM5 -> UGA)	0.1139	[0.1094,0.1184]
			I (SEA -> UGA -> CAM6)	<b>0.5419</b>	<b>[0.5351,0.5486]</b>
			II	0.3442	[0.3379,0.3505]
			III	0.0000	[0.0000,0.0000]
PHL	NGA	UGA	I (SEA -> CAM5 -> UGA)	0.0277	[0.0234,0.0321]
			I (SEA -> UGA -> NGA)	<b>0.7781</b>	<b>[0.7570,0.7991]</b>
			II	0.1942	[0.1746,0.2138]
			III	0.0000	[0.0000,0.0000]
PHL	COG	UGA	I (SEA -> COG -> UGA)	0.1534	[0.2268,0.2419]
			I (SEA -> UGA -> COG)	0.2344	[0.1470,0.1599]
			II	<b>0.6122</b>	<b>[0.6036,0.6208]</b>
			III	0.0000	[0.0000,0.0000]
PHL	CIV	UGA	I (SEA -> CIV -> UGA)	0.0956	[0.0907,0.1005]
			I (SEA -> UGA -> CIV)	0.4244	[0.4161,0.4328]
			II	<b>0.4800</b>	<b>[0.4717,0.4882]</b>
			III	0.0000	[0.0000,0.0000]
PHL	GAB1	COG	I (SEA -> GAB1 -> COG)	0.0941	[0.0875,0.1006]
			I (SEA -> COG -> GAB1)	0.2382	[0.2278,0.2485]
			II	<b>0.6664</b>	<b>[0.6552,0.6776]</b>
			III	0.0014	[0.0010,0.0018]
PHL	GAB2	COG	I (SEA -> GAB2 -> COG)	0.1581	[0.1526,0.1637]
			I (SEA -> COG -> GAB2)	0.1497	[0.1445,0.1550]
			II	<b>0.6906</b>	<b>[0.6839,0.6974]</b>
			III	0.0015	[0.0011,0.0019]
PHL	CAM5	COG	I (SEA -> CAM5 -> COG)	0.1361	[0.1310,0.1413]
			I (SEA -> COG -> CAM5)	0.2169	[0.2110,0.2229]
			II	<b>0.6469</b>	<b>[0.6400,0.6537]</b>
			III	0.0001	[0.0000,0.0001]
PHL	CAM6	COG	I (SEA -> CAM6 -> COG)	0.1385	[0.1326,0.1444]
			I (SEA -> COG -> CAM6)	0.2193	[0.2123,0.2264]
			II	<b>0.6421</b>	<b>[0.6341,0.6501]</b>
			III	0.0001	[0.0000,0.0002]
PHL	NGA	COG	I (SEA -> NGA -> COG)	0.1082	[0.1036,0.1127]
			I (SEA -> COG -> NGA)	0.2702	[0.2640,0.2765]
			II	<b>0.6214</b>	<b>[0.6145,0.6282]</b>
			III	0.0003	[0.0001,0.0004]
PHL	CIV	COG	I (SEA -> CIV -> COG)	0.2344	[0.2281,0.2408]
			I (SEA -> COG -> CIV)	0.1046	[0.0999,0.1092]
			II	<b>0.6604</b>	<b>[0.6533,0.6674]</b>
			III	0.0006	[0.0004,0.0009]
PHL	GAB1	CIV	I (SEA -> GAB1 -> CIV)	0.0843	[0.0794,0.0891]
			I (SEA -> CIV -> GAB1)	0.1774	[0.1706,0.1842]
			II	<b>0.7284</b>	<b>[0.7205,0.7364]</b>
			III	0.0099	[0.0084,0.0115]
PHL	GAB2	CIV	I (SEA -> GAB1 -> CIV)	0.2977	[0.2895,0.3059]
			I (SEA -> CIV -> GAB1)	0.1104	[0.1048,0.1159]
			II	<b>0.5919</b>	<b>[0.5831,0.6006]</b>
			III	0.0001	[0.0000,0.0002]
PHL	CAM5	CIV	I (SEA -> CAM5 -> CIV)	0.4005	[0.3749,0.4261]
			I (SEA -> CIV -> CAM5)	0.1142	[0.1020,0.1264]
			II	<b>0.4853</b>	<b>[0.4601,0.5105]</b>
			III	0.0000	[0.0000,0.0000]
PHL	CAM6	CIV	I (SEA -> CAM6 -> CIV)	0.2265	[0.2100,0.2431]
			I (SEA -> CIV -> CAM6)	0.1417	[0.1300,0.1535]
			II	<b>0.6317</b>	<b>[0.6128,0.6507]</b>
			III	0.0000	[0.0000,0.0000]
PHL	NGA	CIV	I (SEA -> NGA -> CIV)	0.1372	[0.1318,0.1426]
			I (SEA -> CIV -> NGA)	0.1953	[0.1892,0.2015]
			II	<b>0.6674</b>	<b>[0.6601,0.6746]</b>
			III	0.0002	[0.0001,0.0003]

*Confidence in scenario choice.* The model-checking analysis allowed verifying the good adequacy of the best scenarios to the observed data. The observed values of all the summary statistics used were not discordant with the distribution of values obtained on the simulated datasets. This was verified over all alternative analyses (data not shown).

For this same scenario of introduction, we computed a type I error value of 0.29 and a type II error value of 0.13. Type I error represents the probability of rejecting the best scenario whereas it is the true scenario. Such high value does not put back into question our results because other scenarios pointed by simulated datasets were generally those which described close scenarios of divergence to the best scenario (here introduction via an unsampled populations), but very rarely contradictory histories (here two independent introductions or introduction via GAB1). Type II error represented the probability of accepting the best scenario, whereas it is not the true scenario. Again, scenarios most likely to generate such errors were mainly scenarios the closest to the best scenario, making type I error value acceptable.

#### *Propagation pathways within Africa*

*Posterior probabilities.* Posterior probabilities of scenarios for the 6-scenarios analyses are presented in Table 4, with scenarios I, II or IV getting the highest posterior probabilities depending on sampled populations used alternatively in the analyse. Taken in their whole these results are consistent with the scenarios of divergence already pointed by the 3-scenarios analyses, notably the well-supported direct divergence of WEST-group populations from UGA (modelled in scenarios I, II or IV). Despite the results being often ambiguous, with highest posterior probabilities always inferior to 0.5, further main results emerged from these analyses, notably regarding the CIV position. When using UGA with CIV and the WEST-group, scenarios I modelling a regular geographic propagation was the best supported for UGA/GAB2/CIV, UGA/CAM5/CIV and UGA/CAM6/CIV, revealing that CIV might derive from some WEST-group populations. By contrast, CIV did not derive from GAB1 or NGA. CIV would derive more directly from UGA (scenario II being the best supported). The uncertainty associated with CIV position would suggest that there is no clear pathway of dissemination following a geographical logic from WEST-group until CIV, as hypothesized by historical data. Finally, regarding COG position, scenario IV was most frequently the best scenario when using UGA and the WEST-group, suggesting that COG was rather derived independently from the original unsampled African population than directly from UGA (while

**Table 4: Posterior probabilities of scenarios in the 6-scenarios analyses, for different alternative analyses using different populations as POP2 and POP3.**

The confident interval for each scenario is given into brackets. Scenarios I to VI are exposed in Figure 3. Results are shown for SEA corresponding to PHL (Philippines population). Here UGA (Uganda population) always stands for POP1 (in red). WEST-group populations are indicated in blue. For each analysis, the selected scenario is highlighted in bold.

Pop SEA	POP1: UGA	POP2	POP3	Competing scenarios (*selected scenario highlighted)	Probability of scenario (logistic regression)	Confident intervals
PHL	UGA	GAB1	CIV	I (UGA → GAB1 → CIV)	0.0210	[0.0189, 0.0231]
				I (UGA → CIV → GAB1)	0.0289	[0.0260, 0.0318]
				<b>II</b>	<b>0.4511</b>	<b>[0.4353, 0.4669]</b>
				III	0.1773	[0.1670, 0.1876]
				IV (UGA → GAB1)	0.2448	[0.2308, 0.2587]
				IV (UGA → CIV)	0.0405	[0.0370, 0.0439]
				V (GAB1 → CIV)	0.0014	[0.0012, 0.0016]
				V (CIV → GAB1)	0.0033	[0.0028, 0.0038]
				VI	0.0318	[0.0290, 0.0346]
				<b>I (UGA → GAB2 → CIV)</b>	<b>0.2331</b>	<b>[0.2241, 0.2420]</b>
PHL	UGA	GAB2	CIV	I (UGA → CIV → GAB2)	0.1377	[0.1309, 0.1445]
				II	0.1262	[0.1201, 0.1323]
				III	0.1209	[0.1153, 0.1266]
				IV (UGA → GAB2)	0.0278	[0.0257, 0.0299]
				IV (UGA → CIV)	0.0915	[0.0856, 0.0974]
				V (GAB2 → CIV)	0.0966	[0.0909, 0.1023]
				V (CIV → GAB2)	0.0422	[0.0389, 0.0455]
				VI	0.1240	[0.1174, 0.1306]
				<b>I (UGA → CAM5 → CIV)</b>	<b>0.4129</b>	<b>[0.4047, 0.4210]</b>
				I (UGA → CIV → CAM5)	0.1641	[0.1583, 0.1699]
PHL	UGA	CAM5	CIV	II	0.0702	[0.0666, 0.0738]
				III	0.0958	[0.0915, 0.1001]
				IV (UGA → CAM5)	0.0238	[0.0220, 0.0257]
				IV (UGA → CIV)	0.0148	[0.0136, 0.0160]
				V (CAM5 → CIV)	0.1095	[0.1049, 0.1142]
				V (CIV → CAM5)	0.0438	[0.0411, 0.0465]
				VI	0.0652	[0.0618, 0.0687]
				<b>I (UGA → CAM6 → CIV)</b>	<b>0.3072</b>	<b>[0.2920, 0.3225]</b>
				I (UGA → CIV → CAM6)	0.1407	[0.1314, 0.1500]
				II	0.1697	[0.1592, 0.1802]
PHL	UGA	CAM6	CIV	III	0.1342	[0.1256, 0.1428]
				IV (UGA → CAM6)	0.0522	[0.0473, 0.0571]
				IV (UGA → CIV)	0.0289	[0.0260, 0.0319]
				V (CAM6 → CIV)	0.0607	[0.0551, 0.0663]
				V (CIV → CAM6)	0.0275	[0.0246, 0.0304]
				VI	0.0789	[0.0724, 0.0854]
				<b>I (UGA → NGA → CIV)</b>	<b>0.0942</b>	<b>[0.0876, 0.1009]</b>
				I (UGA → CIV → NGA)	0.0937	[0.0873, 0.1001]
				<b>II</b>	<b>0.3679</b>	<b>[0.3550, 0.3809]</b>
				III	0.1837	[0.1746, 0.1927]
PHL	UGA	NGA	CIV	IV (UGA → NGA)	0.1831	[0.1727, 0.1934]
				IV (UGA → CIV)	0.0184	[0.0168, 0.0200]
				V (NGA → CIV)	0.0077	[0.0067, 0.0086]
				V (CIV → NGA)	0.0134	[0.0119, 0.0149]
				VI	0.0379	[0.0350, 0.0409]
				<b>I (UGA → GAB1 → COG)</b>	<b>0.0161</b>	<b>[0.0143, 0.0179]</b>
				I (UGA → COG → GAB1)	0.0396	[0.0359, 0.0434]
				II	0.2725	[0.2573, 0.2876]
				III	0.1066	[0.0989, 0.1143]
				<b>IV (UGA → GAB1)</b>	<b>0.4887</b>	<b>[0.4684, 0.5089]</b>
PHL	UGA	COG	GAB1	IV (UGA → COG)	0.0164	[0.0148, 0.0181]
				V (GAB1 → COG)	0.0019	[0.0016, 0.0022]
				V (COG → GAB1)	0.0091	[0.0079, 0.0103]
				VI	0.0490	[0.0446, 0.0533]
				<b>I (UGA → GAB2 → COG)</b>	<b>0.0162</b>	<b>[0.0150, 0.0174]</b>
				I (UGA → COG → GAB2)	0.0192	[0.0177, 0.0206]
				<b>II</b>	<b>0.2799</b>	<b>[0.2710, 0.2889]</b>
				III	0.1382	[0.1323, 0.1440]
				IV (UGA → GAB2)	0.2478	[0.2378, 0.2579]
				IV (UGA → COG)	0.1388	[0.1317, 0.1458]
PHL	UGA	COG	GAB2	V (GAB2 → COG)	0.0053	[0.0048, 0.0059]
				V (COG → GAB2)	0.0061	[0.0054, 0.0067]
				VI	0.1485	[0.1415, 0.1555]
				<b>I (UGA → CAM5 → COG)</b>	<b>0.0383</b>	<b>[0.0341, 0.0425]</b>
				I (UGA → COG → CAM5)	0.0303	[0.0266, 0.0340]
				II	0.2639	[0.2463, 0.2815]
				III	0.1391	[0.1278, 0.1504]
				<b>IV (UGA → CAM5)</b>	<b>0.3913</b>	<b>[0.3664, 0.4162]</b>
				IV (UGA → COG)	0.0289	[0.0252, 0.0326]
				V (CAM5 → COG)	0.0069	[0.0058, 0.0080]
PHL	UGA	COG	CAM5	V (COG → CAM5)	0.0093	[0.0076, 0.0110]
				VI	0.0921	[0.0823, 0.1018]
				<b>I (UGA → CAM6 → COG)</b>	<b>0.0855</b>	<b>[0.0798, 0.0912]</b>
				I (UGA → COG → CAM6)	0.0622	[0.0577, 0.0666]
				II	0.2792	[0.2679, 0.2905]
				III	0.1253	[0.1188, 0.1318]
				<b>IV (UGA → CAM6)</b>	<b>0.3190</b>	<b>[0.3041, 0.3339]</b>
				IV (UGA → COG)	0.0331	[0.0303, 0.0359]
				V (CAM6 → COG)	0.0166	[0.0149, 0.0184]
				V (COG → CAM6)	0.0188	[0.0168, 0.0208]
PHL	UGA	COG	CAM6	VI	0.0603	[0.0559, 0.0646]
				<b>I (UGA → NGA → COG)</b>	<b>0.0322</b>	<b>[0.0275, 0.0368]</b>
				I (UGA → COG → NGA)	0.0432	[0.0375, 0.0489]
				II	0.3043	[0.2817, 0.3269]
				III	0.1164	[0.1051, 0.1278]
				<b>IV (UGA → NGA)</b>	<b>0.4632</b>	<b>[0.4339, 0.4924]</b>
				IV (UGA → COG)	0.0104	[0.0088, 0.0121]
				V (NGA → COG)	0.0023	[0.0018, 0.0029]
				V (COG → NGA)	0.0073	[0.0059, 0.0088]
				VI	0.0207	[0.0177, 0.0236]
PHL	UGA	COG	CIV	<b>I (UGA → CIV → COG)</b>	<b>0.0456</b>	<b>[0.0400, 0.0512]</b>
				I (UGA → COG → CIV)	0.0730	[0.0650, 0.0810]
				<b>II</b>	<b>0.2604</b>	<b>[0.2434, 0.2775]</b>
				III	0.1370	[0.1260, 0.1479]
				<b>IV (UGA → CIV)</b>	<b>0.2631</b>	<b>[0.2424, 0.2838]</b>
				IV (UGA → COG)	0.0478	[0.0423, 0.0534]
				V (CIV → COG)	0.0087	[0.0071, 0.0104]
				V (COG → CIV)	0.0192	[0.0161, 0.0223]
				VI	0.1451	[0.1316, 0.1587]



WEST-group populations were founded from UGA). COG remained thus independent from all other populations.

*Confidence in scenario choice.* The model-checking analysis allowed verifying the good adequacy of the best scenarios to the observed data. The observed values of all the summary statistics used were not discordant with the distribution of values obtained on the simulated datasets. This was verified over all alternative analyses (data not shown).

### Parameter estimates

Medians of the posterior distributions were used as point estimates for parameters. Estimations of time parameters were biologically consistent across analyses (data not shown). Parameters estimations for the scenario of divergence of GAB1 from UGA are detailed hereafter. We found here 585 generations before present for the first African introduction, representing almost 60 years before present, and 301 generation for GAB1 divergence representing 30 years before present. Regarding demographic parameters, effective sizes of populations were estimated in the order of magnitude of 70000 for ancestral SEA population and 35000 for African populations. Bottlenecks were characterized by their severity corresponding to the ratio between their duration (*db*) and the number of individuals (*Nb*) (Pascual *et al.* 2007, Lombaert *et al.* 2010). We found here an intensity of 0.53 for UGA, representing 94 individuals during 50 generations, and 0.66 for GAB1 (representing 76 individuals during 50 generations). The results of bottleneck intensities for the different populations were consistent across all alternative analyses, and in the same order of magnitude for all the different African populations considered (data not shown).

High values of RMAE (data not shown) reflected that parameters were not precisely estimated (data not shown). Their order of magnitude deserved nevertheless to be examined because they were coherent and very consistent across the whole analyses of the two rounds.

## **DISCUSSION**

### ***A single original African introduction***

The first aim of this study was to investigate the scenario of introduction of *Mycosphaerella fijiensis* into Africa within a statistical framework. The main hypothesis was that all our samples would derive from the same initial founding population, based on the drastic

**Table S1: Posterior distribution of parameters for the 3-scenarios analysis considering the divergence of GAB1 from UGA (Scenario I in the analysis including PHL, GAB1 and UGA, Table 3).**

Parameter	Mean	Median	Mode	Quantile 2.5%	Quantile 97.5%
$t_{GAB1}$	3.44E+02	3.01E+02	1.82E+02	1.57E+02	7.46E+02
$t_{UGA}$	6.11E+02	5.85E+02	4.58E+02	3.27E+02	9.70E+02
$t_{PHL}$	7.40E+03	7.85E+03	9.52E+03	2.64E+03	9.91E+03
$N_{GAB1}$	4.10E+04	3.67E+04	4.86E+02	1.22E+03	9.58E+04
$N_{UGA}$	4.03E+04	3.46E+04	4.73E+03	1.97E+03	9.61E+04
$N_{PHL}$	6.70E+04	6.80E+04	7.07E+04	3.05E+04	9.70E+04
$db_{GAB1}$	3.11E+01	3.30E+01	5.01E+01	3.47E+00	4.98E+01
$Nb_{GAB1}$	7.72E+01	4.41E+01	1.77E+01	5.40E+00	3.71E+02
$db_{UGA}$	3.19E+01	3.44E+01	4.94E+01	3.61E+00	5.00E+01
$Nb_{UGA}$	1.18E+02	5.01E+01	2.25E+01	6.90E+00	4.64E+02
$\mu_{mic}$	2.01E-05	1.55E-05	7.07E-06	2.92E-06	6.51E-05
$t_{GAB1} * \mu_{mic}$	6.50E-03	5.00E-03	2.67E-03	9.00E-04	2.06E-02
$t_{UGA} * \mu_{mic}$	1.17E-02	9.22E-03	5.76E-03	1.88E-03	3.59E-02
$t_{PHL} * \mu_{mic}$	1.45E-01	1.13E-01	6.37E-02	2.06E-02	4.63E-01
$N_{GAB1} * \mu_{mic}$	8.32E-01	4.97E-01	1.40E-02	1.50E-02	3.72E+00
$N_{UGA} * \mu_{mic}$	7.53E-01	4.61E-01	5.90E-02	3.10E-02	3.28E+00
$N_{PHL} * \mu_{mic}$	1.19E+00	1.03E+00	7.56E-01	2.56E-01	3.13E+00
$db_{GAB1} * \mu_{mic}$	6.31E-04	4.51E-04	1.46E-04	3.30E-05	2.30E-03
$Ng_{GAB1} * \mu_{mic}$	1.54E-03	6.80E-04	7.00E-05	5.00E-05	8.33E-03
$db_{UGA} * \mu_{mic}$	6.55E-04	4.64E-04	1.34E-04	3.40E-05	2.38E-03
$Ng_{UGA} * \mu_{mic}$	2.15E-03	8.80E-04	1.50E-04	8.00E-05	1.19E-02

$t_{GAB1}$ ,  $t_{UGA}$ ,  $t_{PHL}$  : divergence time of GAB1, UGA and South-East Asian population, respectively (in number of generations before present);

$N_{GAB1}$ ,  $N_{UGA}$ ,  $N_{PHL}$  : effective size of GAB1, UGA and PHL populations, respectively;

$db_{GAB1}$ ,  $db_{UGA}$  : bottleneck duration for GAB1 and UGA populations, respectively (in number of generations);

$Nb_{GAB1}$ ,  $Nb_{UGA}$  : effective size during bottleneck for GAB1 and UGA populations, respectively;

$\mu_{mic}$  : microsatellite mutation rate.

reduction of genetic diversity within Africa compared to South-East Asia, and the relative homogeneity in allelic composition across African populations (Robert *et al.* 2012). This has been formally confirmed here, with strong support, by systematically rejecting scenario of two separate introductions on both sides of the continent, whatever the pair of sampled populations used.

Moreover the place where original introduction occurred has been further uncovered here. According to the most commonly expressed historical hypothesis invasion of (at least) West Africa would have resulted from a Gabonese introduction of infected plant material brought from Asia, in 1978 (Frossard 1980, Mourichon & Fullerton 1990). An independent introduction in African east coast was moreover suspected to have led to spread of the disease across East Africa. The unofficial observation of the disease in Zambia in 1973 (Raemaekers 1978) had also suggested that a more central introduction could have occurred earlier than in Gabon. Descriptive population genetics analyses were not consistent with a Gabonese introduction. They suggested rather a more ancient establishment of Uganda population, based on its higher genetic diversity and its substantial private allelic richness compared to other African populations, including Gabon populations. The scenario of introduction in Gabon was definitely rejected by ABC analyses, and this for both populations sampled in Gabon GAB1 and GAB2. It remains in fact possible that an introduction of infected plant material in Gabon might have lead to a local epidemic, but not to invasion of the whole continent. The quality control of our analyses (type II error and model checking) and the number of replicate analyses run with different populations allow to be confident in this interpretation. If initial introduction had taken place in Gabon, we would expect that some of the geographically neighboring populations also appear among the earliest established populations. By contrast, ABC tests gave the best support to the scenarios where GAB1, CAM5, CAM6 or NGA diverged from UGA, and little support to the reverse scenario.

Can we then suspect that the original African introduction took place in Uganda? We can argue here that UGA population is more closely derived from the initial founding population of Africa than any other of our sampled populations, but the results of our ABC analyses are actually not clear-cut enough to infer more precisely the place of introduction event. The scenario of introduction via UGA was not systematically the highest probability scenario, sometimes supplanted by scenario of introduction via an unsampled population. As we did not dispose of an exhaustive sampling, interpretations of these Bayesian analyses have to remain careful: the best supported scenario by the data is not inevitably the true scenario. More



eastern samples would then be necessary to investigate the historical ideas of central (near Zambia) or eastern introductions of the disease.

### ***Stochastic propagation of the disease within the continent***

At the intra-continental level, the historical hypothesis was that the disease could have progressed regularly, according to a geographical logic, from Gabon to Ivory Coast via Cameroun and Nigeria (Frossard 1980, Mourichon & Fullerton 1990). The second round of ABC analyses helped us to explore this issue. It seems that UGA but also COG would have derived independently from the unknown original African founding population close to UGA. COG then appears rather isolated from the rest of sampled populations. Most of the western populations would then have derived from UGA as supported in previous ABC analysis. The question now is to assess pathways of spread from the WEST-group populations to CIV. Regarding the historical divergence of Ivory Coast population, the whole analyses supported that CIV was always the last founded population, consistently with the description of late disease arrival in this country (Mouliom Pefoura *et al.* 1996). However, according to historical hypotheses, we should expect that CIV systematically diverge directly from WEST-group populations in all ABC tests. By contrast analyses with the different populations of the WEST-group were divided, and CIV position remains unclear in relation to these populations. Notably CIV would be more directly derived from UGA than from GAB1 or NGA, this latter being still precisely located on the pathway toward CIV. Such results could reflect unordered jumps of the disease rather than a gradual geographical pathway from Gabon to Ivory Coast. Such an interpretation is also supported by the fact high levels of differentiation, and slightly different histories, can be observed between close populations of the WEST-group, and even between different populations of the same country. For instance the inclusion of GAB1 or GAB2 in ABC analyses did not always lead to the same scenario of divergence with other populations.

Such a stochastic pattern of spread seems more complex than expected, but could be finally strikingly concordant with the population genetic studies conducted at finer scales in Africa and America (Halkett *et al.* 2010, Rieux *et al.* 2011). Colonization pathways of *M. fijiensis* at the scale of banana producing areas of several hundreds kilometres was suggested to combine both a gradual spread by regular disease front resulting from wind-dispersed ascospores, and genetic signatures of more stochastic processes. Especially, in Cameroon, the existence of a sharp genetic discordance in a banana producing area of several hundred kilometres wide was



shown (Halkett *et al.* 2010, Rieux *et al.* 2011). No geographical or physical barrier could explain such genetic discordance, and one of the plausible explanations would be the occurrence of stochastic founder events through rare natural long-distance dispersal event, or infected plant material exchanges. Our study support that both regular and stochastic mechanisms were indeed implicated in the spread of the disease at the scale of the African continent.

Such a pattern of stochastic continental propagation, creating important levels of genetic structure at all scales, appears remarkable for a wind-dispersed pathogen spreading on large uniform banana plantations. It notably contrasts with the low population structure and regular gene flow highlighted in wheat pathogens *Mycosphaerella graminicola* or *Phaeosphaeria nodorum* (Banke & McDonald 2005, Stukenbrock *et al.* 2006) and apple pathogen *Venturia inaequalis* (Gladieux *et al.* 2008). A possible explanation is that human dissemination of *M. fijiensis* can only occur via infected leaves but not fruits, limiting the impact of man on gene flow. A more exhaustive sampling and further landscape genetic studies would thus be necessary to further understand the relative contributions of progression by gradual front or by stochastic jumps, for such a rapid invasion of a whole continent.

### ***Demographical variations in invading populations***

Despite a lack of precision, orders of magnitude of demographical parameters estimated with ABC methods were consistent across alternative analyses, providing important pieces of information that deserved to be discussed. Firstly, as expected, very high effective size of several tens thousands individuals were estimated for South-east Asian and African populations. Such values appear much higher than the averages found in other animal or plant organism. *M. fijiensis* populations sampled in this study mainly come from large industrial banana plantations, which could partly account for such results. This is comparable with Zhan & McDonald (2004) estimations in *M. graminicola* on wheat supporting that the particular biology of such fungal organisms is probably responsible for these huge effective sizes.

Compared to this, we highlighted that populations also experienced drastic bottlenecks, estimated in average by the presence of only a few tens of individuals during a period of a few tens of generations. Such estimations probably reflect that each sampled population underwent several successive founder events during its history within Africa. Indeed we showed that such events occurred not only at the entrance into the continent, but also associated with subsequent spread of the disease across the continent. Our results reveal rather





consistent with analyses performed with the MCMC-based method MsVar (Beaumont 1999, Storz & Beaumont 2002) by Rieux (2011) on a Cameroonian population, also estimating a drastic bottleneck under a different demographic model (i.e., instantaneous reduction of effective size, immediately followed by exponential growth of the population until present). We further confronted results of to these two different methods, by computing the total “genetic drift intensity” undergone by Cameroonian populations (by using CAM6 in our analysis). In the two cases, the total intensity of genetic drift was thus integrated over the period going from the initial African founder event until present, this time taking into account all population effective size variations estimated during this time period (in our analyses CAM6 underwent a first bottleneck at the initial African introduction via UGA, followed by recovered high population effective size, and then a second bottleneck at its own foundation followed by high recovered effective size). Our analyses lead to a “genetic drift intensity” of 7.6, representing 66 effective individuals during a period fixed to 500 generations (i.e., 50 years), while we found 2.4 from Rieux (2011) MsVar analyses, representing around 210 effective individuals during 500 generations. Drift intensity is thus lower in MsVar estimates than in the present ABC ones, although of the same order of magnitude. All these concordant results highlight anyway drastic bottlenecks generating high genetic drift intensity, probably resulting from the introduction of few propagules.

Bottlenecks are in fact frequently suggested in fungal plant pathogens invasions, generally at the entrance into a continent after long distance dispersal events (Stukenbrock *et al.* 2007, Gladieux *et al.* 2008, Barres *et al.* 2008, Linde *et al.* 2009, Zaffarano *et al.* 2009). By contrast such strong continental genetic structure resulting from widespread occurrence of founder events has rarely been shown at the intra-continental scale, except for the barley pathogen *Rhynchosporium secalis* (Linde *et al.* 2009). This study is one of the first that provides estimations of bottlenecks associated with fungal pathogens invasions, at both inter and intra-continental level.

### ***General conclusion***

In conclusion, the present study allows to progress towards understanding of the different patterns of invasions followed by pathogens of cultivated plant. We showed here, with statistical tests and demographic parameters values estimations that a fungal pathogen of banana leaves was able to invade a whole continent from a single drastic founder event, followed by other multiple bottlenecks associated with intra-continental spread. Our results



should interestingly be compared with further examples, to assess if such drastic demographic fluctuations are general features among emerging fungal plant pathogens. Here the resulting loss of neutral genetic diversity obviously did not prevent *M. fijiensis* from emerging and invading entire continents, which raises further general questions regarding the importance of molecular genetic diversity for plant pathogen invasion success.



## REFERENCES

- Anderson PK, Cunningham AA, Patel NG, *et al.* (2004) Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution* **19**, 535-544.
- Banke S, McDonald BA (2005) Migration patterns among global populations of the pathogenic fungus *Mycosphaerella graminicola*. *Molecular Ecology* **14**, 1881-1896.
- Barres B, Halkett F, Dutech C, *et al.* (2008) Genetic structure of the poplar rust fungus *Melampsora larici-populina*: Evidence for isolation by distance in Europe and recent founder effects overseas. *Infection Genetics and Evolution* **8**, 577-587.
- Beaumont MA (1999) Detecting population expansion and decline using microsatellites. *Genetics* **153**, 2013-2029.
- Beaumont MA, Zhang WY, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics* **162**, 2025-2035.
- Bertorelle G, Benazzo A, Mona S (2010) ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular Ecology* **19**, 2609-2625.
- Carlier J, Lebrun MH, Zapater MF, Dubois C, Mourichon X (1996) Genetic structure of the global population of banana black leaf streak fungus, *Mycosphaerella fijiensis*. *Molecular Ecology* **5**, 499-510.
- Chakraborty R, Jin L (1993) A unified approach to study hypervariable polymorphisms: statistical considerations of determining relatedness and population distances. *Exs* **67**, 153-175.
- Churchill ACL (2010) *Mycosphaerella fijiensis*, the black leaf streak pathogen of banana: progress towards understanding pathogen biology and detection, disease development, and the challenges of control. *Molecular Plant Pathology* **12**, 307-328.
- Cornuet J-M, Ravigne V, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *Bmc Bioinformatics* **11**.
- Cornuet J-M, Santos F, Beaumont MA, *et al.* (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* **24**, 2713-2719.
- Desprez-Loustau ML, Robin C, Buee M, *et al.* (2007) The fungal dimension of biological invasions. *Trends in Ecology & Evolution* **22**, 472-480.
- Engelbrecht CJB, Harrington TC, Steimel J, Capretti P (2004) Genetic variation in eastern North American and putatively introduced populations of *Ceratocystis fimbriata* f. *platani*. *Molecular Ecology* **13**, 2995-3005.
- Estoup A, Guillemaud T (2010) Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology* **19**, 4113-4130.
- Facon B, Genton BJ, Shykoff J, *et al.* (2006) A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology & Evolution* **21**, 130-135.
- Frossard P (1980) Apparition d'une nouvelle et grave maladie foliaire des bananiers et plantains au Gabon : la maladie des raies noires : *Mycosphaerella fijiensis* Morelet. *Fruits* **35**, 519-527.
- Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology* **10**, 305-318.
- Gauhl F (1994) Epidemiology and ecology of black sigatoka (PhD Thesis), INIBAP, Montpellier
- Gladieux P, Zhang X-G, Afoufa-Bastien D, *et al.* (2008) On the Origin and Spread of the Scab Disease of Apple: Out of Central Asia. *Plos One* **3**.

- Goldstein DB, Linares AR, Cavallisforza LL, Feldman MW (1995) An evaluation of genetic distances for use with microsatellite loci. *Genetics* **139**, 463-471.
- Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity* **86**, 485-486.
- Guillemaud T, Beaumont MA, Ciosi M, Cornuet JM, Estoup A (2010) Inferring introduction routes of invasive species using approximate Bayesian computation on microsatellite data. *Heredity* **104**, 88-99.
- Halkett F, Coste D, Platero GGR, *et al.* (2010) Genetic discontinuities and disequilibria in recently established populations of the plant pathogenic fungus *Mycosphaerella fijiensis*. *Molecular Ecology* **19**, 3909-3923.
- Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403-1405.
- Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *Bmc Genetics* **11**, 94.
- Jones DR (2000) Introduction to banana, abaca and enset. In: Diseases of Banana, Abaca and Enset (ed. Jones DR), pp. 37-79. CABI Publishing, Wallingford, UK.
- Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes* **5**, 187-189.
- Linde CC, Zala M, McDonald BA (2009) Molecular evidence for recent founder populations and human-mediated migration in the barley scald pathogen *Rhynchosporium secalis*. *Molecular Phylogenetics and Evolution* **51**, 454-464.
- Lombaert E, Guillemaud T, Cornuet J-M, *et al.* (2010) Bridgehead Effect in the Worldwide Invasion of the Biocontrol Harlequin Ladybird. *Plos One* **5**.
- Montarry J, Andrivon D, Glais I, *et al.* (2010) Microsatellite markers reveal two admixed genetic groups and an ongoing displacement within the French population of the invasive plant pathogen *Phytophthora infestans*. *Molecular Ecology* **19**, 1965-1977.
- Mouliom-Pefoura A, Lassoudière A, Foko J, Fontem DA (1996) Comparison of development of *Mycosphaerella fijiensis* and *Mycosphaerella musicola* on banana and plantain in the various ecological zones in Cameroon. *The American Phytopathological Society* **80**, 950-954.
- Mourichon X, Fullerton RA (1990) Geographical distribution of the two species *Mycosphaerella musicola* Leach (*Cercospora musae*) and *M. fijiensis* Morelet (*C. fijiensis*), respectively agents of Sigatoka disease and Black Leaf Streak disease in bananas and plantains. *Fruits* **45**, 213-218.
- Munkacsı AB, Stoxen S, May G (2008) Ustilago maydis populations tracked maize through domestication and cultivation in the Americas. *Proceedings of the Royal Society B-Biological Sciences* **275**, 1037-1046.
- Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89, 583-590.
- Pascual M, Chapuis MP, Mestres F, *et al.* (2007) Introduction history of *Drosophila subobscura* in the New World: a microsatellite-based survey using ABC methods. *Molecular Ecology* **16**, 3069-3083.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.
- Raboin LM, Selvi A, Oliveira KM, *et al.* (2007) Evidence for the dispersal of a unique lineage

- from Asia to America and Africa in the sugarcane fungal pathogen *Ustilago scitaminea*. *Fungal Genetics and Biology* **44**, 64-76.
- Raymond M, Rousset F (1995) Genepop (version-1.2) - population-genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.
- Rieux A. (2011) Etude des processus de dispersion et des flux géniques chez un champignon phytopathogène : le cas de *Mycosphaerella fijiensis* à l'échelle d'un bassin de production Camerounais. PhD thesis, Montpellier SupAgro
- Rieux A, Halkett F, de Bellaire LD, *et al.* (2011) Inferences on pathogenic fungus population structures from microsatellite data: new insights from spatial genetics approaches. *Molecular Ecology* **20**, 1661-1674.
- Rivas GG, Zapater MF, Abadie C, Carlier J (2004) Founder effects and stochastic dispersal at the continental scale of the fungal pathogen of bananas *Mycosphaerella fijiensis*. *Molecular Ecology* **13**, 471-482.
- Rousset F (2008) GENEPOP '007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**, 103-106.
- Storey JD (2002) A direct approach to false discovery rates. *Journal of the Royal Statistical Society Series B-Statistical Methodology* **64**, 479-498.
- Stukenbrock EH, Banke S, Javan-Nikkhah M, McDonald BA (2007) Origin and domestication of the fungal wheat pathogen *Mycosphaerella graminicola* via sympatric speciation. *Molecular Biology and Evolution* **24**, 398-411.
- Stukenbrock EH, Banke S, McDonald BA (2006) Global migration patterns in the fungal wheat pathogen *Phaeosphaeria nodorum*. *Molecular Ecology* **15**, 2895-2904.
- Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.
- Zaffarano PL, McDonald BA, Linde CC (2008) Rapid speciation following recent host shifts in the plant pathogenic fungus *Rhynchosporium*. *Evolution* **62**, 1418-1436.
- Zaffarano PL, McDonald BA, Linde CC (2009) Phylogeographical analyses reveal global migration patterns of the barley scald pathogen *Rhynchosporium secalis*. *Molecular Ecology* **18**, 279-293.
- Zhan J, McDonald BA (2004) The interaction among evolutionary forces in the pathogenic fungus *Mycosphaerella graminicola*. *Fungal Genetics and Biology* **41**, 590-599.
- Zhou X, Burgess TI, De Beer ZW, *et al.* (2007) High intercontinental migration rates and population admixture in the sapstain fungus *Ophiostoma ips*. *Molecular Ecology* **16**, 89-99.

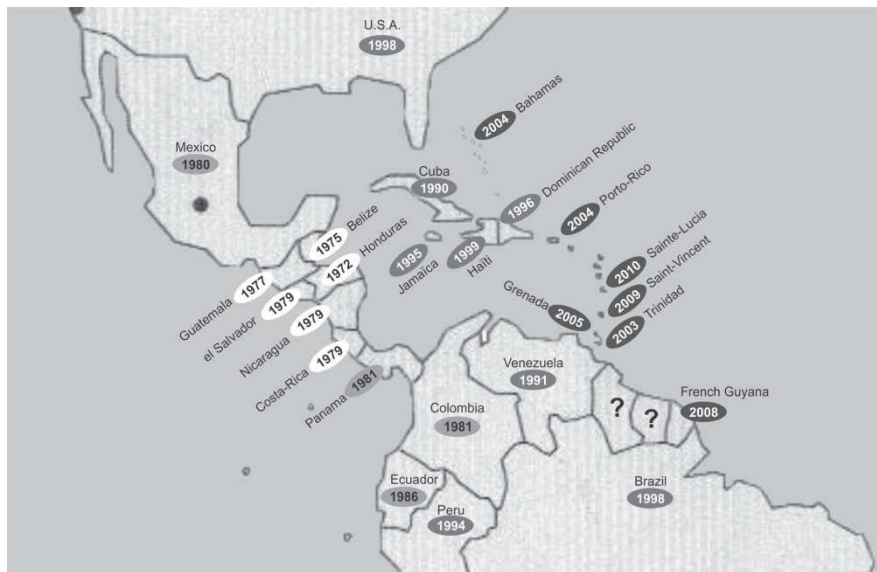




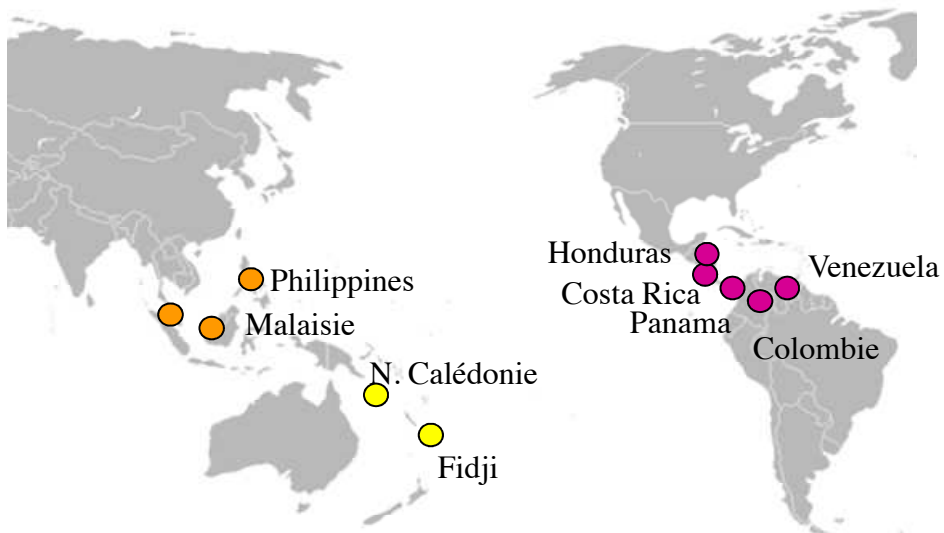
## **Histoire de l'invasion de *M. fijiensis* en Amérique : processus d'introduction et d'établissement en contextes continental et insulaire.**

Le contexte et les problématiques fondamentales de cette étude rejoignent celles de l'étude Africaine, cherchant à mieux comprendre les processus d'invasion d'une maladie fongique de cultures à l'échelle d'un continent. Dans notre première étude à échelle mondiale (Chapitre I), nous suggérions, sur la base d'analyses de réseaux de séquences et de classification bayésienne sur microsatellites, que l'introduction de la maladie sur le continent Américain pourrait avoir été accompagnée d'un événement de mélange entre des populations sources de différentes origines (admixture), probablement de l'Océanie et du Sud-Est Asiatique. Nous montrions aussi une tendance (non significative) à une diminution de la diversité génétique le long d'un front continental allant du Honduras jusqu'au Venezuela. Par ailleurs, Halkett *et al.* (2010) sur un transect de 300km au Costa Rica, suggéraient également la propagation d'un front de colonisation régulier qui aurait laissé une trace sous forme d'isolement par la distance à cette échelle. Tous ces indices semblent indiquer que les populations issues d'un mélange initial au Honduras se seraient répandues suivant une progression géographique associée à des événements de fondation successifs plus ou moins sévères. Dans cette étude, le premier objectif est dans un premier temps de tester dans un cadre probabiliste cette hypothèse. D'autre part, cette étude a pour but d'aller plus loin dans la compréhension de la propagation au sein de l'Amérique, en y incluant cette fois-ci l'archipel des Caraïbes, zone où la maladie progressait encore récemment (Figure 1). Peut-on caractériser la propagation récente au sein des Caraïbes ? Les analyses génétiques confirment-elles les voies et modes de dissémination supposés par les données de suivi historique ? Que peuvent-t-elles nous apprendre quant aux événements démographiques subis par les populations aux cours de ces sauts d'île en île ?

Deux originalités se dessinent ici par rapport à l'étude conduite au sein du continent africain. La première, de nature fondamentale, est l'opportunité de reconstruire et mieux comprendre la propagation d'une telle maladie fongique au sein d'un système qui inclut cette fois-ci une forte composante insulaire, supposant forcément le franchissement de distances géographiques de longueur variables sans aucune présence de bananier. La deuxième originalité est de nature appliquée, puisque une meilleure connaissance des processus de dissémination et d'établissement de la maladie au sein de l'archipel des Caraïbes peut fournir



**Figure 1 :** Progression de la maladie des raies noires en Amérique : année de première détection au sein des pays (d'après Lapeyre de Bellaire *et al.* 2010).



**Figure 2 :** Localisation géographique des populations échantillonnées, utilisées dans l'étude ABC.

des informations précieuses en matière de prévention (bien que la grande majorité des Caraïbes soit déjà infectée) mais également de contrôle.

Nous avons pris le parti de présenter séparément ici les analyses répondant aux deux grands objectifs de cette étude. Premièrement nous présentons une analyse ABC approfondie dont l'objectif était de tester l'hypothèse de mélange de populations à l'introduction sur le continent, et d'inférer les paramètres démographiques associés à cet événement. Deuxièmement nous présentons une analyse de génétique des populations plus classique, réalisée au sein de l'équipe, à laquelle j'ai participé sans la mener directement, afin de proposer une première description des patrons de diversité et de structure des populations dans les Caraïbes. Cette analyse reste donc très préliminaire. L'ensemble de ces études devra être finalement intégré, mais nécessitera pour cela un approfondissement des analyses concernant la propagation intra-continentale, et notamment au niveau des Caraïbes.

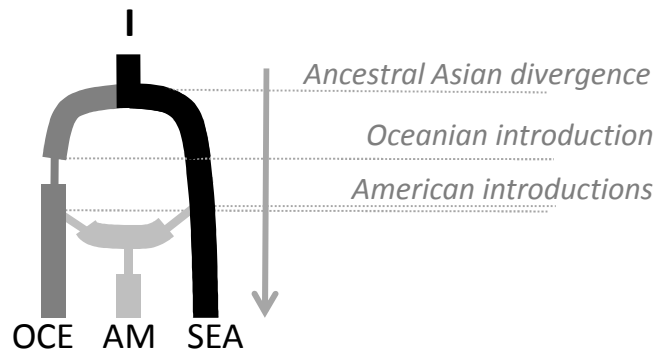
### **Introduction de *M. fijiensis* en Amérique : test du scénario de mélange de populations**

*Auteurs: S. ROBERT, M.-F. ZAPATER, J. CARLIER et V. RAVIGNE*

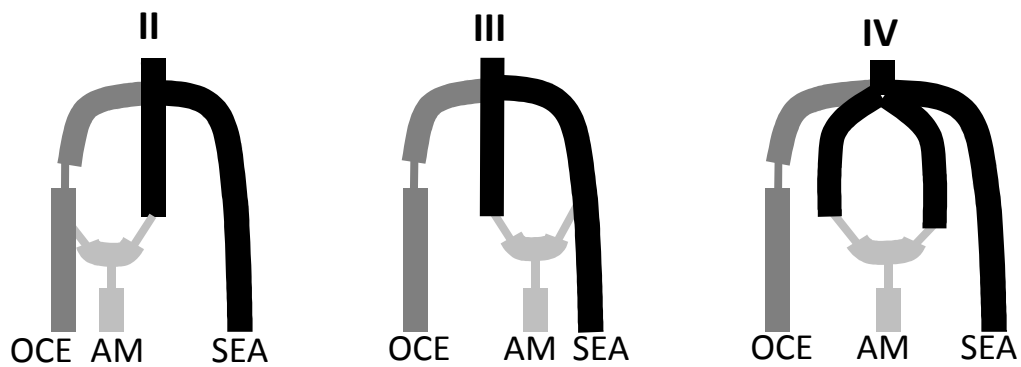
#### **Contexte : données historiques et étude génétique mondiale**

Les données historiques suggèrent fortement que l'introduction de la maladie ait eu lieu au Honduras, quelques années avant sa première détection en 1972, mais serait restée « quiescente » avant le déclenchement des premières épidémies (Stover & Dickson 1976, Stover 1980). Celles-ci ont débuté au sein et à proximité d'une collection de diverses variétés de bananiers et plantains en provenance de différents pays d'Asie du Sud Est, rassemblée dans les années 1960 (Rosales *et al.* 1999). Comme suggéré dans l'étude du Chapitre 1, l'hypothèse d'une admixture entre sources différentes pourrait parfaitement bien être expliquée par ces transports de bananiers. Cependant, une autre source insoupçonnée en provenance d'Océanie est aussi fortement suggérée par cette étude. Enfin, données historique comme étude mondiale du Chapitre 1 suggèrent que la maladie aurait ensuite été disséminée (naturellement et/ou par l'homme) de façon graduelle de pays en pays jusqu'au Venezuela (Mourichon & Fullerton 1990).

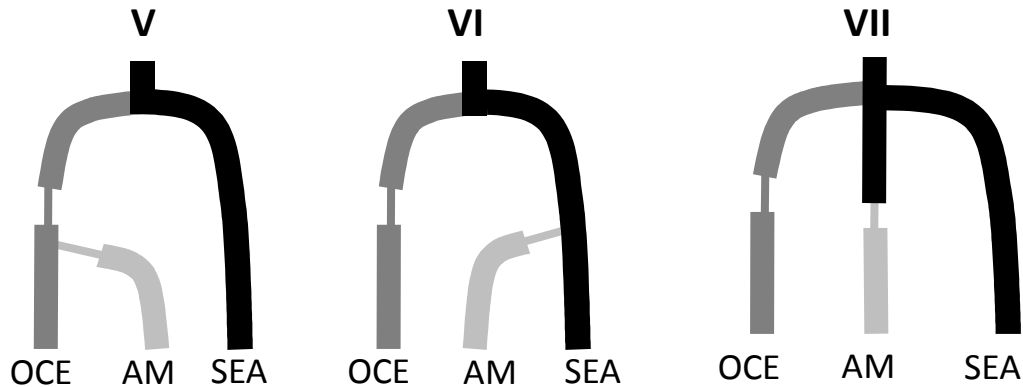
## Scénario imaginé



## Variantes plausibles



## Hypothèses alternatives



**Figure 3 :** Scénarios en compétition dans l'analyse ABC. SEA: population échantillonnée Sud-Est Asiatique, OCE: population échantillonnée d'Océanie, AM: population échantillonnée américaine. L'épaisseur des traits matérialise les variations potentielles de tailles efficaces des populations au moment des introductions (goulots d'étranglement).

## Matériel et méthodes

Des populations de l'étude mondiale en provenance d'Asie du Sud Est, d'Océanie et d'Amérique, génotypées via le panel mondial de 21 marqueurs microsatellites de Robert *et al.* (2012), ont été réinvesties ici (Figure 2). Nous avons employé une procédure ABC avec le logiciel DIYABC (Cornuet *et al.* 2008, 2010) comme détaillé dans l'étude Africaine, en considérant 7 scénarios en compétition (Figure 3). Ces scénarios incluent 3 populations échantillonnées en provenance du Sud-Est Asiatique (SEA), d'Océanie (OCE) et d'Amérique (AM). Le scénario I représente l'hypothèse, soutenue par l'étude mondiale du Chapitre I, selon laquelle la population américaine dériverait d'un mélange entre une population d'Océanie et d'Asie du Sud Est. Les scénarios II, III et IV modélisent des variantes de ce scénario, en y introduisant la possibilité que les véritables populations sources (asiatique et océanienne) de l'introduction américaine ne soient pas directement disponibles comme échantillons dans cette étude. On considère donc les possibilités de mélange entre OCE et une population non échantillonnée (scénario II), entre SEA et une population non échantillonnée (scénario III), et entre deux populations non échantillonnées (scénario IV). Enfin les scénarios V, VI et VII représentent des hypothèses alternatives sans mélange : l'introduction aurait lieu de manière directe à partir de OCE (scénario V), de SEA (scénario VI), ou enfin d'une population non échantillonnée (scénario VII). Les tests ont été répliqués en utilisant différentes populations en tant que SEA (PHL, MAL1 et MAL2), OCE (FIJ et Ncl) et AM (HDN, CR, PAN, COL, VEN), afin de vérifier la constance des résultats.

## Résultats et Discussion

### ***Scénario d'introduction en Amérique et paramètres démographiques***

Après vérification, nous présentons ici uniquement les résultats des tests effectués avec les populations des Philippines pour le Sud-Est Asiatique (SEA), et de Fidji pour l'Océanie (OCE) (Table 1). Dans toutes les analyses alternatives, le scénario II reçoit le meilleur support (avec des probabilités postérieures allant de 0.50 à 0.74, erreur de type I de 0.70 et erreur de type II de 0.088), supportant donc fortement l'existence d'un évènement d'admixture entre OCE et une population non échantillonnée (que l'on peut supposer dans le Sud-Est Asiatique). Notons que deux autres scénarios obtiennent un certain support dans tous les tests : le scénario I d'admixture directe entre SEA et OCE, et le scénario V de divergence directe de la population américaine à partir de OCE. Nous pouvons dire que l'ensemble de ces

**Table 1:** Probabilités postérieures des scénarios, pour les différents tests effectués entre alternant les populations américaines. Dans ces tests la population du Sud Est Asiatique SEA correspond à PHL et la population d'Océanie OCE correspond à FIJ.

American population (AM)	Competing scenarios (selected scenario highlighted)	Probability of scenario (logistic regression)
<b>H</b> (Honduras)	I	0.1760 [0.1596,0.1923]
	<b>II</b>	<b>0.6378 [0.6114,0.6642]</b>
	III	0.0001 [0.0001,0.0001]
	IV	0.0005 [0.0004,0.0006]
	V	0.1852 [0.1622,0.2082]
	VI	0.0000 [0.0000,0.0000]
	VII	0.0004 [0.0003,0.0006]
<b>H-G</b> (Honduras)	I	0.1245 [0.1106,0.1385]
	<b>II</b>	<b>0.7432 [0.7194,0.7671]</b>
	III	0.0010 [0.0008,0.0013]
	IV	0.0069 [0.0055,0.0083]
	V	0.1183 [0.0989,0.1378]
	VI	0.0000 [0.0000,0.0001]
	VII	0.0059 [0.0047,0.0072]
<b>CRI</b> (Costa Rica)	I	0.2101 [0.1930,0.2273]
	<b>II</b>	<b>0.5823 [0.5569,0.6077]</b>
	III	0.0001 [0.0000,0.0001]
	IV	0.0004 [0.0003,0.0005]
	V	0.2066 [0.1839,0.2294]
	VI	0.0000 [0.0000,0.0000]
	VII	0.0004 [0.0003,0.0005]
<b>PAN</b> (Panama)	I	0.2217 [0.2059,0.2374]
	<b>II</b>	<b>0.5095 [0.4865,0.5324]</b>
	III	0.0000 [0.0000,0.0000]
	IV	0.0001 [0.0000,0.0001]
	V	0.2688 [0.2466,0.2909]
	VI	0.0000 [0.0000,0.0000]
	VII	0.0001 [0.0000,0.0001]
<b>COL</b> (Colombie)	I	0.1693 [0.1539,0.1846]
	<b>II</b>	<b>0.5740 [0.5465,0.6015]</b>
	III	0.0005 [0.0004,0.0006]
	IV	0.0031 [0.0024,0.0037]
	V	0.2507 [0.2230,0.2783]
	VI	0.0000 [0.0000,0.0000]
	VII	0.0025 [0.0019,0.0030]
<b>VEN</b> (Venezuela)	I	0.2515 [0.2291,0.2739]
	<b>II</b>	<b>0.6058 [0.5790,0.6327]</b>
	III	0.0048 [0.0038,0.0058]
	IV	0.0168 [0.0138,0.0199]
	V	0.1049 [0.0884,0.1213]
	VI	0.0005 [0.0003,0.0006]
	VII	0.0157 [0.0125,0.0189]

résultats supporte bien l'histoire d'admixture caractérisant l'introduction américaine, et soutient particulièrement l'existence d'au moins une source en provenance d'Océanie.

Les estimations de paramètres sous le scénario sélectionné (II), bien que peu précises, sont très constantes d'un test à un autre et méritent d'être examinées (tableau non présenté). Les tailles efficaces de populations, de l'ordre de 70000 pour SEA et de l'ordre de 40000 pour OCE et AM sont parfaitement concordantes avec les résultats obtenus à partir des analyses africaines, indiquant là encore des valeurs bien supérieures aux ordres de grandeurs caractérisant d'autres types d'organismes. Le taux de mélange sont estimés entre 0.4 et 0.5, ce qui appuie l'idée que la source OCE n'a pas contribué seule à la constitution des populations américaines. La sévérité du goulot d'étranglement associé à la fondation de la population américaine est estimée à environ 0.4, soit 125 individus sur 50 générations, donc là encore de l'ordre de grandeur des estimations de goulots d'étranglement en Afrique (bien qu'à peine moins drastique). Ceci montre que la population issue de mélange initiale a pu subir un fort goulot d'étranglement. Ce résultats ne nous permettent pas d'inférer la manière dont se sont déroulées les introductions, mais on pourrait supposer qu'il y ait eu une introduction (ou plusieurs au même endroit) associée à la création de la collection de bananiers (introduisant potentiellement déjà des individus différenciés si différents bananiers infectés ont été collectés) et une autre introduction à partir d'Océanie. Cette dernière serait particulièrement plausible au vu des fréquents mouvements de matériel végétal en provenance du Pacifique, via Hawaï, décrits par Stover (1978) à cette époque. Les données historiques rapportent l'existence d'un délai au Honduras, avant le déclenchement des premières grandes épidémies en 1972, car des indices suggèrent que la maladie aurait été présente au sein de la collection de bananiers depuis la fin des années 60 (Stover & Dickson 1976). Ainsi ce délai pourrait correspondre soit à un effet purement démographique (effet Allee) ainsi qu'aux variations climatiques, soit au temps nécessaire à une adaptation des populations.

Finalement, toutes les populations du continent américain (H et H-G, mais également CR, PAN, COL et VEN) soutiennent le même scénario d'admixture (II), ce qui supporte l'hypothèse (historique et issue de notre étude mondiale), selon laquelle la maladie se serait ensuite propagée jusqu'au Venezuela à partir des populations issues de mélange initiales du Honduras.

### ***Remarques sur la progression continentale***

**Table 2:** Fst calculés par paires de populations (tiré de Robert *et al.* 2012)

	COL	CR	H	H-G	MEX	PAN
COL						
CR	0,33					
H	0,30	0,02				
H-G	0,19	0,17	0,13			
MEX	0,41	0,17	0,12	0,23		
PAN	0,30	0,07	0,10	0,20	0,23	
VEN	0,34	0,37	0,32	0,21	0,47	0,42



Un tableau des Fst par paires de population, issu de l'analyse mondiale, est présenté à nouveau ici (Table 2), afin de se focaliser les patrons de structure au sein du continent. On observe notamment que les niveaux de différenciation entre paires de populations, relativement faibles entre populations proches (par exemple Honduras/Costa Rica) ont tendance à augmenter avec la distance géographique, ce qui va de concert avec la diminution de diversité intra-population déjà relevée dans notre étude précédente. Aucune de ces observations ne s'est avérée significative statistiquement. Par ailleurs, l'interprétation de ces patrons de structure peut être compliquée par l'écart entre les dates d'échantillonnage des populations de 1999 et de 2009. Cependant ces tendances suggèrent une propagation très graduelle de la maladie, où le front de progression des ascospores pourrait jouer un rôle important. On trouve en effet dans ces régions de grands bassins de production bananière susceptibles de couvrir des surfaces presque continues sur plusieurs centaines de kilomètres, comme au Costa Rica (Halkett *et al.* 2010). Ceci n'exclut pas l'existence d'évènements de fondation au cours de la progression, et on suggère qu'un important goulot d'étranglement aurait ainsi eu lieu à l'introduction au Venezuela, probablement depuis la Colombie (fort Fst Colombie/Venezuela et faible diversité du Venezuela, Robert *et al.* 2012). Cela serait particulièrement concordant avec les caractéristiques de paysage et les barrières montagneuses que le pathogène aurait eu à franchir entre les deux pays, et ce de manière plus probable par un transport humain que par une dissémination des spores.

### **Chemins d'expansion de la maladie dans les Caraïbes**

*Auteurs: M.-F. ZAPATER, S. ROBERT, V. ROUSSEL, L. PIGNOLET, J. CARLIER et C. ABADIE*

#### **Données et hypothèses historiques**

Une synthèse des données historiques de détection et suivi spatio-temporel de la maladie au sein de chaque île, a d'autre part été menée par Catherine Abadie. Ces données, plus précises dans les Caraïbes que dans le reste du monde, devrait être de plus fiables indicateurs de l'arrivée de la maladie, voire de ses routes de dispersion au sein des îles pour les mieux renseignées.

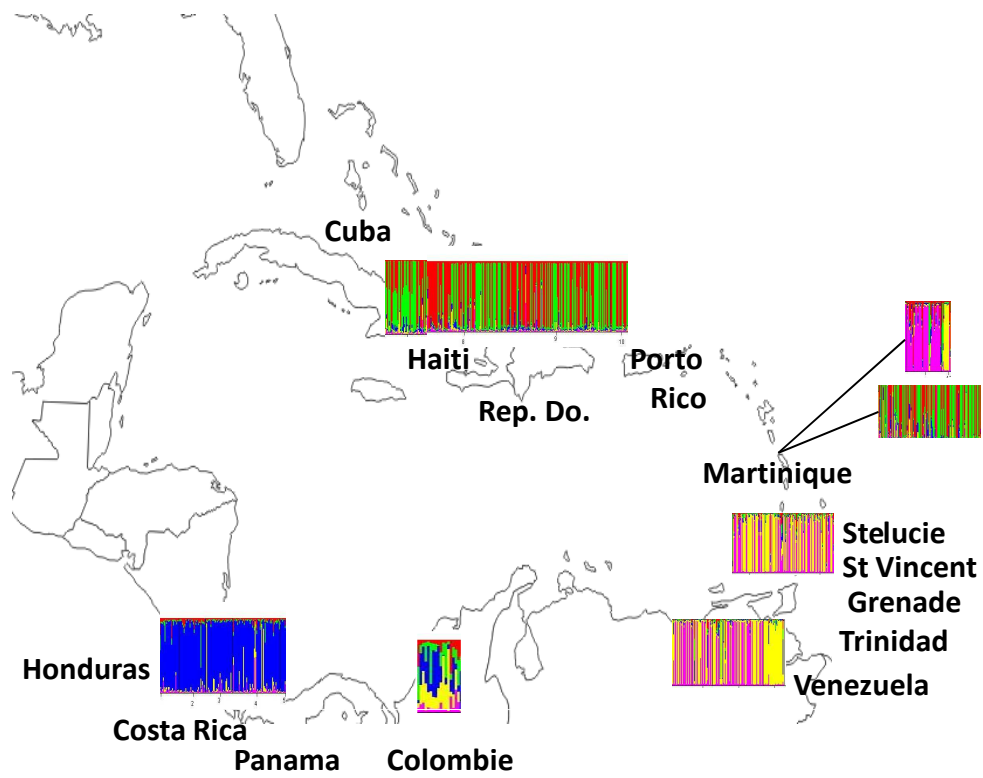


La maladie n'a cessé de progresser au sein de l'archipel depuis sa première introduction à Cuba en 1990 (Jones 2000, Figure 1). Malgré les mesures préventives mises en place par les pays menacés, elle a progressé vers l'Est d'île en île, mais à une vitesse particulièrement faible selon Pasberg-Gauhl *et al.* (2000), pour qui les vents d'Est dominants auraient ralenti la progression naturelle de la maladie par le vent. Récemment, un autre front de propagation supposé venir du continent a permis l'arrivée de la maladie sur les îles proches du Venezuela, de Trinidad à Sainte Lucie. Les Antilles françaises ainsi encerclées faisaient partie des dernières îles indemnes, mais malgré les mesures préventives de quarantaine, *M. Fijiensis* a finalement été détecté pour la première fois en Martinique à l'automne 2010. Au vu des prospections précises qui ont suivi à travers l'île, la maladie aurait été présente depuis un certain temps (quelques semaines à quelques mois), couvrant déjà une importante surface de l'île. Compte tenu des mesures phytosanitaires en place dans les Antilles, l'hypothèse principale avançant qu'une dispersion naturelle d'ascospores par le vent, en provenance des îles proches de Saint Vincent et Sainte Lucie, serait responsable de cette introduction.

### Matériel et méthodes

Les nouvelles populations analysées dans cette étude sont issues d'envois d'échantillons foliaires par des partenaires ou collectés dans l'équipe, à partir desquels l'extraction d'ADN a été réalisée principalement sur des isolats, et plus rarement des lésions foliaires (Robert *et al.* 2010). De nouvelles populations réparties dans l'archipel des Caraïbes (Cuba, Jamaïque, Haïti, République Dominicaine, Porto Rico, Martinique, Sainte Lucie, Saint Vincent, Grenade, Trinidad), ainsi que deux nouvelles populations en Colombie et au Venezuela, ont ainsi pu être génotypées avec un nouveau panel de 18 marqueurs microsatellites adaptés pour des études à l'échelle de l'Amérique. Cela représente un échantillon de 3 à 4 populations par île en moyenne, afin de couvrir au maximum l'archipel. Nous avons effectué des analyses descriptives préliminaires sur cet échantillon total, notamment afin d'explorer la répartition et structure de la diversité génétique, et tenter de déceler d'éventuelles sous-structures au sein des îles. Nous avons conduit également une analyse de classification bayésienne avec STRUCTURE (Pritchard *et al.* 2000), (analyses non présentées).

D'autres analyses descriptives et de classification bayésienne ont été réalisées sur les 14 microsatellites en commun entre le panel de l'étude mondiale, et celui de l'étude Caraïbes, afin d'y inclure à la fois les nouvelles populations des Caraïbes, et celles du continent utilisées



**Figure 4 :** Résultats de l'analyse préliminaire de clustering effectuée avec STRUCTURE (K=5).

dans notre précédente étude. En se basant sur les résultats de l'analyse Caraïbes, nous avons alors conservé un nombre réduit de populations par île dès lors que les différentes populations d'une même île s'avéraient faire partie du même cluster génétique.

### Résultats et discussion

Les analyses présentées en Figure 4 ont été réalisées sur les 14 microsatellites en commun entre le panel de l'étude mondiale, et celui de l'étude Caraïbe. On y retrouve les observations décrites ci-dessus, avec une discontinuité génétique entre la Colombie et le Venezuela (qui supporte notre hypothèse préalable de fondation depuis la Colombie puisque les individus colombiens sont malgré tout assignés de façon minoritaire au groupe génétique du Venezuela). Les résultats des analyses ont montré de plus de façon très nette que deux groupes génétiques différents couvrent majoritairement les Caraïbes. En liaison avec les données historiques, on aurait ainsi affaire à deux fronts de propagation différenciés, l'un en provenance du Venezuela et couvrant les îles en direction du Nord, et l'autre dans la partie Ouest des Caraïbes, provenant probablement des populations du continent bien que les analyses présentées nous renseignent peu sur ce point. Concernant les pays pour lesquels nous disposons de plusieurs populations couvrant l'île, la seule sous-structure détectée se trouve précisément en Martinique, où les deux fronts d'avancée de la maladie se rejoignent, donnant lieu notamment à une population mélangée échantillonnée vers le centre de la Martinique.

Le cas de la Martinique mérite que l'on s'y arrête plus précisément. De manière étonnante, le groupe génétique correspondant au front Sud (englobant les populations du Venezuela jusqu'à Sainte Lucie) se retrouve dans la partie Nord de la Martinique, tandis que le front correspondant aux îles Nord-Ouest des Caraïbes est retrouvé dans la partie Sud de la Martinique, incluant notamment les zones proches du port et de l'aéroport, où la maladie était présente depuis un certain temps (Abadie, comm. pers.). Nous avons affaire en Martinique à un contact secondaire entre deux fronts de propagation de la maladie, résultant probablement d'au moins deux introductions différentes. Un tel constat était loin d'être soupçonné *a priori*. Si ces analyses génétiques préliminaires ne nous permettent pas de démontrer les causes naturelles vs. anthropiques de ces introductions, de forts soupçons pèsent néanmoins sur une introduction par l'homme (voire les deux), supposant probablement une enfreinte aux mesures de préventions.



## Perspectives

Ces résultats devront être confirmés avec des analyses d'assignement plus approfondies, notamment des tests ancrés dans les statistiques bayésiennes, prenant en compte les délais d'échantillonnage entre différentes populations. L'emploi des approches de génétique du paysage pourrait également s'avérer précieuses à partir d'échantillonnages 2D réalisés au sein des îles (Archie *et al.* 2009, Biek & Real 2010). Ces méthodes pourraient d'ailleurs être mises en œuvre au sein de chacune des îles et intégrées avec les données de suivi de maladie, afin de mieux comprendre la dynamique de dispersion de *M. fijiensis* à faible échelle sur des aires délimitées que sont les îles. De plus, des données retraçant éventuellement des mouvements humains ou des épisodes climatiques tels que les cyclones susceptibles de transporter des fragments de matériel végétal, s'avéreraient naturellement très précieuses.

Enfin, si l'on peut montrer que le mélange de groupes génétiques observé au sein d'une population en Martinique est effectif, d'intéressantes perspectives d'études théoriques pourraient s'ouvrir en Martinique, visant notamment à explorer le rôle potentiel du mélange de populations dans le succès d'invasion de ce champignon phytopathogène.





# CHAPITRE 3



## CHAPITRE 3

### **Comparaison des stratégies de vie des cercosporioses du bananier : un rôle dans le remplacement par *Mycosphaerella fijiensis* de *M. musicola* ?**

Lors de son invasion mondiale, *M. fijiensis* n'est pas arrivé dans un environnement vierge, mais dans des agrosystèmes au sein desquels l'espèce proche *M. musicola* était déjà établie. L'émergence mondiale de *M. musicola* a précédé dans le temps celle de *M. fijiensis*, et les données historiques relatant son suivi mondial suggèrent qu'elle proviendrait également d'Asie du Sud-Est et du Pacifique avant d'avoir été disséminée sur les autres continents dans les années 1930-1960 (Jones 2000). *M. fijiensis* et *M. musicola* ont des cycles de vie très semblables et génèrent des symptômes très proches sur les feuilles de bananier. Or, lors de son invasion mondiale, *M. fijiensis* a supplanté *M. musicola* en l'espace de quelques années, dans toutes ses régions d'introduction (hormis dans quelques productions d'altitude au climat plus frais, peu tolérées par *M. fijiensis*, Jones 2000). De plus *M. fijiensis* a partout été décrit comme causant des épidémies plus sévères et se propageant plus rapidement. L'objet de ce chapitre est d'explorer quels traits d'histoire de vie clés pourraient conférer son succès d'établissement à *M. fijiensis*, et notamment sa supériorité compétitive sur *M. musicola*.

Pour cela nous voulons comparer les deux espèces dans un même environnement. Nous avons mesuré des traits d'histoire de vie liés à la partie asexuée du cycle en conditions contrôlées, ainsi qu'à partir d'échantillons foliaires collectés au champ en Martinique par Catherine Abadie et Yolande Chilin-Charles (BGPI), zone où la présence simultanée des deux espèces est encore observable en raison de l'arrivée très récente de *M. fijiensis*. D'importantes mises au point ont préalablement été nécessaires (menées avec Marie-Françoise Zapater), et la partie sexuée du cycle reste encore trop complexe à étudier dans ces conditions de laboratoires. Quant aux échantillons foliaires collectés, aucune plage nécrotique liée à la reproduction sexuée de *M. fijiensis* n'a été trouvée sur les bananiers considérés, où *M. fijiensis* s'est avéré encore très minoritaire par rapport à *M. musicola*. Enfin signalons que les analyses statistiques menées sur les valeurs de traits d'histoire de vie restent très préliminaires, basées la plupart du temps sur des tests non paramétriques simples, et devront être approfondies dans la mesure du possible.



Par ailleurs un modèle épidémiologique dérivé des modèles SI, développé par Virginie Ravigné, Valérie Lemesle (CIRAD – UR 26) et Catherine Abadie (BGPI), nous a de plus servi de base de réflexion théorique préliminaire pour adresser les questions posées, et a donc été intégré à l'étude présentée ici. Ce modèle a l'originalité de prendre en compte explicitement les deux types de spores « sexuées » et « asexuées » de ces espèces, mais devra être à terme spatialisé et étendu à la compétition afin de pouvoir apporter des éléments de réponse théorique aux questions posées.

L'étude dans son ensemble est présentée sous la forme d'un article non publié.



# **Asexual life strategies differ between two related fungal pathogens in competition on the same plant.**

**The case of *Mycosphaerella fijiensis* and *M. musicola* on banana.**

S. ROBERT<sup>\*</sup>, C. ABADIE<sup>\*</sup>, M.-F. ZAPATER<sup>\*</sup>, F. CARREEL<sup>\*</sup>, L. PIGNOLET<sup>\*</sup>, STEEVE JOSEPH<sup>\*</sup>, Y. CHILIN-CHARLES<sup>\*</sup>, R. HABAS<sup>\*</sup>, J. CARLIER<sup>\*</sup> and V. RAVIGNE<sup>\*</sup>

<sup>\*</sup> CIRAD, UMR BGPI, Campus international de Baillarguet, TA A-54K, F-34398 Montpellier Cedex 5, France

## **ABSTRACT**

Successive invasions and dominance of increasingly problematic plant pathogens within agro-ecosystems is a crucial issue in plant epidemiology. Identifying key life-history traits that determine the competitive success of new strains or species over resident ones is thus an important issue, remaining rarely investigated. Here we addressed this issue with a pathosystem of two close fungal pathogen species in competition on banana hosts. The ascomycete *Mycosphaerella fijiensis* causing the Black Leaf Streak Disease of banana has recently spread worldwide and rapidly supplanted its previously established sister species, *M. musicola*, responsible for the Sigatoka disease. To better understand the role of life-history traits expressed in their complex life cycle, we developed a new epidemiological model suitable for both species, taking into account both sexual and asexual reproduction. Based on this latter, we underlined the substantial role of the asexual part of the life cycle for the establishment and spread in new environments, even in such sexually reproducing pathogens. We then compared asexual life history strategies between the two species, under controlled conditions in laboratory and in field conditions. We show that asexual reproduction occurs earlier but generates much less spores in *M. fijiensis* than in *M. musicola*, revealing potential trade-offs between timing and capacities of asexual reproduction. We discuss potential implications of such life strategies for the fitness advantage of *M. fijiensis* over *M. musicola*.





## INTRODUCTION

With the acceleration of global exchanges, multiple introductions of invasive or pathogenic organisms have become frequent (Anderson *et al.* 2004, Estoup & Guillemaud 2010, Hufbauer *et al.* 2012). In agriculture, the risks of pandemics of large-scale crops caused by novel pathogen strains, and emergences resulting from introductions of exotic pathogens, are expected to increase (Stukenbrock & McDonald 2008). Recent epidemics provide examples of such pathogens, for example the Dutch elm disease in northwest Europe (Brasier 2001) or the black pod rot of cocoa in Africa (Acebo-Guerrero *et al.* 2012). New strains of *Phytophthora infestans*, the pathogenic oomycete that is known to have caused Irish famine of the mid-XIX<sup>th</sup> century, have also recently emerged in Europe causing again devastating epidemics on potato (Cooke *et al.* 2011). A dramatic epidemic of wheat stem rust caused by race UG99 of the fungus *Puccinia graminis tritici* is currently spreading worldwide (Singh *et al.* 2011). In many instances (such as the latter two), the emergent pathogen spreads on hosts already attacked by other more or less related pathogens. Successive strain replacement is even a major mode of evolution in asexual plant pathogen populations, such as wheat rust *Puccinia sp.* (Goyeau *et al.* 2007, Bahri *et al.* 2009, Hovmoller *et al.* 2011). This is why identifying factors that control invasion and persistence of new crop pathogens and their ability to outcompete resident strains are among the main issues of epidemiology (Gilligan & van den Bosch 2008, Gilligan 2008).

To do so, one needs to understand how the life history strategy of a new pathogenic strain confers a relative fitness advantage over a resident strain, i.e., a higher effective contribution to the next generations. Life-history traits directly condition the important events of the life cycles of species linked to growth, reproduction and death, and particularly regulate the timing of reproduction and number of offspring. According to life history theory, life history strategies are adjusted by different correlations and trade-offs between traits depending on the limited level of resources available to invest in different traits (Stearns 1992, Roff 1992, Reznick *et al.* 2000). Examples of serial replacements in invasive species have for example shown that life-history traits of new invaders tended to become more and more typical of the so-called K strategies (such as lower progeny size and higher offspring survival), displaying thus higher and higher competitive abilities than previous residents. Such pattern has been well documented for freshwater snails in Martinique island (Facon *et al.* 2006, 2008) or fruit flies in La Réunion island (Duyck *et al.* 2007).



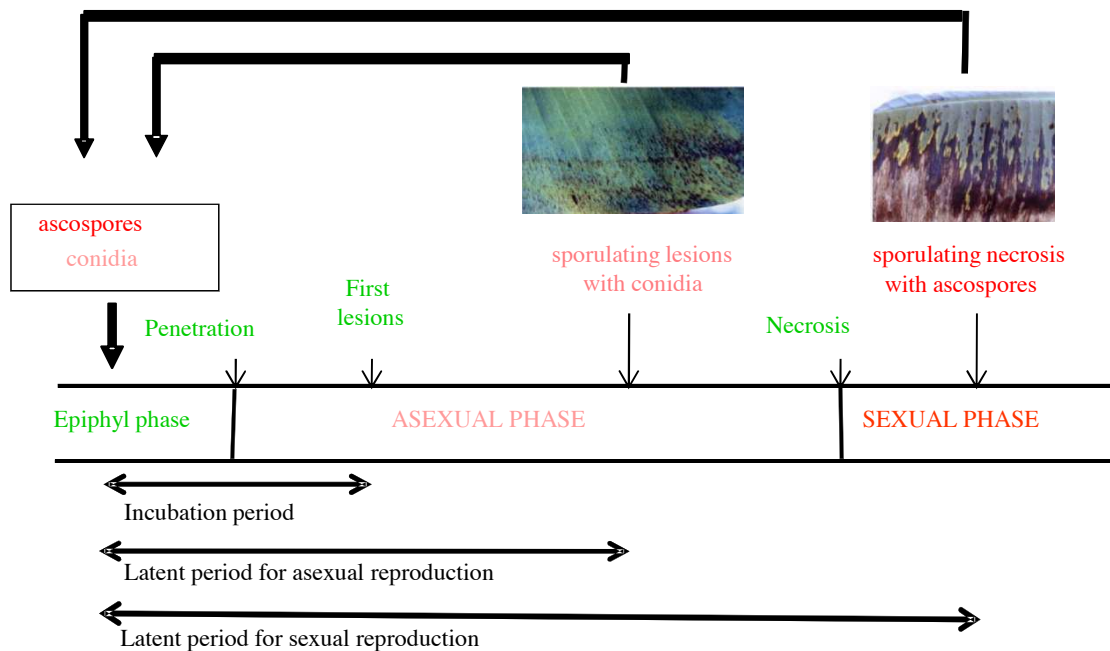
At present it is not known whether such findings could be transposed to fungal plant pathogens, the most important taxonomic group responsible for emerging plant diseases (after viruses, Desprez-Loustau *et al.* 2007). One reason might be that the great diversity of fungal life cycles and their propensity to create multiple kinds of sexual and asexual propagules renders the determination of the most appropriate fitness metrics difficult (Pringle & Taylor 2002). A consequence is that fitness is often measured through life-history traits which effect on fitness is not trivial (Pariaud *et al.* 2009a). Common fitness-related traits are the length of latency period (time interval between infection and the onset of sporulation), spore production capacities or length of infectious period (time during which sporulation occurs). Most of these fitness-related life-history traits correspond in fact to the aggressiveness components generally measured in phytopathological studies. (Miller *et al.* 1998, Hovmoller *et al.* 2008, Milus *et al.* 2008, Pariaud *et al.* 2009a). Authors have generally found that the new dominant fungal strains displayed more efficient values concerning all life history traits considered compared to the resident strains. For instance, the newly dominant strains of wheat leaf rust (*Puccinia triticina*) exhibited higher lesion expansion rate and earlier and more important spore production than former strains (Milus & Line 1980, Miller *et al.* 1998, Milus *et al.* 2008, Pariaud *et al.* 2009b). Interestingly these life-history strategies shifts seem not to reveal trade-offs between life history traits measured here, contrary to what can be observed in animals or plants (Stearns 1992, Roff 1992, Iwasa & Levin 1995).

Understanding how different life-history traits combine to influence fitness represents a particular challenge in fungal pathogens, raising the need for theoretical studies. The most common framework is the family of SI models, that study spatial and temporal dynamics of an epidemic by modeling the flow of host tissue through a succession of classes from healthy (susceptible S) through infected (I) and so to dead tissue (Gilligan & van den Bosch 2008). Life-history traits related to spore production, latent period and infectious period are then considered as determining the rate of flows between the different classes (Vallavieille-pope *et al.* 2000). Criteria for invasion are derived from these models in order to predict invasion and persistence of pathogens, depending of these biologically meaningful parameters. As the widely used  $R_0$  (basic reproductive number), these criteria allow epidemiological insight into why some epidemics take off and others do not, and why some strains or pathotypes come to dominate pathogen populations (Gilligan & van den Bosch 2008). But few existing models really take into account complex life strategies of fungi, and only Gilchrist *et al.* (2006) derived a fitness measure in function of fungal pathogen traits. To our knowledge, only one



type of spore is always modelled while many fungal pathogens produce multiple sexual and asexual spores during their life cycle.

The case of the Black Leaf Streak Disease of bananas (BLSD) caused by the ascomycete *Mycosphaerella fijiensis*, provides an original system to investigate how an introduced pathogen is able to supplant a close competitor species that was previously established. *M. fijiensis* was discovered in Fiji in 1963 (Rhodes 1964) and spread to all banana producing areas across the world in the last 50 years, being still progressing in the Caribbean archipelago (Jones 2000, Marin *et al.* 2003, Churchill 2010). In a few years this global invasion has led to the exclusion of the sister species *Mycosphaerella musicola*, the causal agent of the Sigatoka disease (SD). SD had previously spread and established in all banana plantations worldwide from its first mention in Java at the beginning of the XXth century (Zimmermann 1902). BLSD and SD have the same basic life cycle (Jones 2000) including the production of both asexual and sexual spores (conidia and ascospores respectively). They cause similar foliar symptoms developing from lesions to necrosis of banana leaves, resulting in banana yield losses (Marin *et al.* 2003). BLSD has become the most important foliar pathogen of bananas, revealing more difficult to control than SD under tropical climate. BLSD epidemics have been reported everywhere to develop faster than SD and to cause more severe damages in banana plantations (Jones 2000). Few experimental studies measuring life history traits have been carried out on these fungi. These mainly aimed at comparing different cultivar or environmental conditions rather than different strains. Very few different strains were thus included, or if many strains, authors measured a global disease severity without decomposing life history traits (Mouliom Pefoura & Mourichon 1990, Jacome & Schuh 1993, Fullerton & Olsen 1995, Romero & Sutton 1997). Phytopathological descriptions based on field observations have all underlined the higher ascospore production and a lower conidia production in *M. fijiensis* in comparison to *M. musicola* (Stover 1980, Gauhl 1994). However literature is generally not clear-cut and sometimes contradictory regarding the precise balance of asexual and sexual reproduction in the two species. Ascospores are commonly suspected to play a preponderant role for establishment, long-distance spread and competitive superiority of *M. fijiensis* (Meredith 1970, Stover 1980, Gauhl 1994, Fouré *et al.* 1984), but the relative importance of conidia production is less studied (Burt 1997, Jones 2000, Marin *et al.* 2003). For instance, while Jacome *et al.* (1994) suggested that conidia can be responsible for substantial amount of diseases, their potential importance is rather minimized in other studies (Fouré *et al.* 1984, Marin *et al.* 2003).



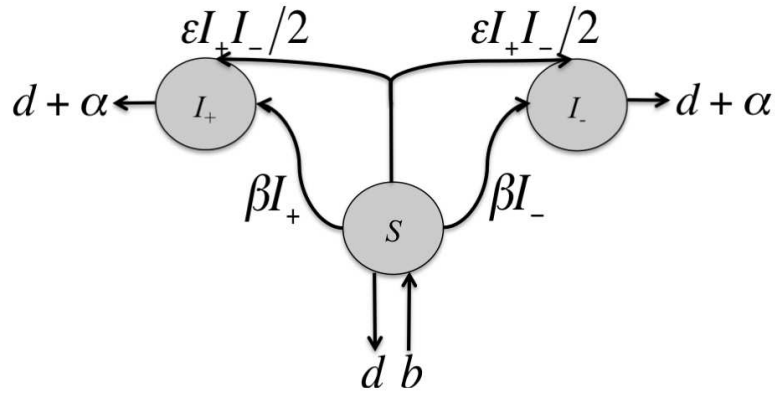
**Figure 1: Basic life cycle shared by the two species, with succession of asexual and sexual spore production.**

Here we propose to examine whether life-history differences between the two species can be detected and whether they may account for the replacement of *M. musicola* by *M. fijiensis*. First, because it is not a trivial task to establish the link between fungal fitness and life-history traits, we develop an epidemiological model deriving from SI models, that considers both asexual and sexual reproduction. We show that the asexual part of the life cycle is critically important to the establishment and spread into new environments. Consequently, we examine and properly compare the combination of life-history traits linked to the asexual part of the life cycle between the two species in common environments. This has been done in two different contexts: firstly in laboratory by controlling artificial inoculations and climatic conditions, and secondly in a natural infestation situation where transitory cohabitation of the two species could be observed. We show that asexual spore production is much more limited and occurs earlier in *M. fijiensis*. Possible implications for the fitness advantage of *M. fijiensis* and exclusion of *M. musicola* are discussed.

## MATERIALS & METHODS

### Species life cycles

The two species have a common hemibiotrophic basic life cycle (Jones 2000, Figure 1). After spore germination and penetration of stomata, mycelium grows within banana leaf tissues resulting in visible growing lesions. Asexual spores (conidia) are firstly produced at lesions surface by the development of reproduction structures in leaf stomata. *M. fijiensis* produces few geniculate conidiophores per stomata, carrying 3 to 4 conidia, whereas *M. musicola* produces dense fascicles of conidiophores called sporodochia, with each conidiophore carrying a single conidia. Several asexual cycles can occur and re-infest older leaves from sporulating lesions on younger leaves, before the pathogen enters in the sexual part of its cycle, with coalescence of leaf lesions and apparition of necrotic areas on banana leaves. In such coalescent lesions, sexual production of spores (ascospores) may occur between lesions of compatible sexual sign (the two species are heterothallic). Conidia are mainly dispersed by rain splash to neighbor leaves and plants over a few meters. Wind-dispersed ascospores can be transported over at least several tens of meters in *M. musicola* (Brun 1963, Stover 1980), and several hundreds of meters in *M. fijiensis* (Burt 1997, Rieux 2011, Abadie & Bonnot *in*



**Figure 2: Diagram of the model, describing the transitions between the densities of healthy – or susceptible - leaf patches (S), and infectious leaf patches (I).** The healthy host population turn-over is characterized by a natality rate  $b$  and a basic mortality rate  $d$ . On the one hand, healthy host patches can become infectious after infection by asexually produced spores (conidia) with sexual sign + or - , at a rate  $\beta$ . On the other hand, healthy host patches can become infectious after infection by sexually produced spores (ascospores) with sexual sign + or - , at a rate  $\varepsilon$ . Contrary to conidia that are directly produced by infectious  $I_+$  or  $I_-$  patches, ascospores production requires the encounter of two compatible patches  $I_+$  and  $I_-$ . Finally infectious patches experience an increased mortality rate  $\alpha$ , corresponding to virulence of the disease.



*prep.*) and are thought to allow long distance dispersal. Ascospores predominantly infect the youngest receptive cigar leaves.

### **Modelling epidemiology with both sexual and asexual spores**

We developed a model derived from classical SI models, taking into account both asexual (conidia) and sexual (ascospores) spores production. Model assumptions are summed up in Figure 2. We consider a population of host plant surface units with constant natality ( $b$ ) and a density-dependent mortality (with rate  $d$ ). Area units of the host population are here represented by banana foliar patches of the mean size of a lesion. So we do not consider growth of lesions. In the absence of any disease, host dynamics is given by:

$$\frac{dS}{dt} = b - dS, \quad (1)$$

where  $S$  is the density of healthy plant area units.

At demographic equilibrium, host population size is:

$$S^* = \frac{b}{d} \quad (2)$$

We further consider a disease that is able to propagate through two types of spores: one is produced asexually and the other is produced sexually when two lesions of compatible sexual signs meet. Let  $I_+$  be the density of leaf surface units that is infected by pathogens with sexual sign + and  $I_-$  the density of leaf surface units that is infected by pathogens with sexual sign -. Surface units become infected by asexual spores at a rate  $\beta$  and by sexual spores at a rate  $\varepsilon$ . Infected surface units experience an increased mortality rate  $\alpha$  that corresponds to the disease virulence. The spatial arrangement of lesions and the dispersal distances of spores are neglected. The dynamics of healthy and infected leaf surface units is then governed by the follow set of equations (Figure 2):

**Table 1: Geographical origin and sampling date of the 19 *M. musicola* strains and 25 *M. fijiensis* strains used for artificial inoculation experiments.**

Species	Strain number	Strain name	Country	Sampling date
<i>M. musicola</i>	1	0061CMR	Cameroon	1987
	2	CB10	Cameroon	1999
	3	CA38	Cameroon	1999
	4	CA39	Cameroon	1999
	5	0062CMR	Cameroon	1987
	6	0038CIV	Ivory Coast	1986
	7	0493GNI	Guinea	1994
	8	0490GIN	Guinea	1994
	9	0127COL	Colombia	1989
	10	0257COL	Colombia	1991
	11	0125COL	Colombia	1989
	12	0168COL	Colombia	1990
	13	328GLP	Guadeloupe	1992
	14	130VEN	Venezuela	1989
	15	500IDN	Indonesia	1988
	16	0913IND	Indonesia	1988
	17	B1.1 a	Indonesia	2010
	18	B 3.1 a	Indonesia	2010
	19	0920IND	Indonesia	1988
<i>M. fijiensis</i>	20	89CMR	Cameroon	
	21	CAM 05-03	Cameroon	2001
	22	COM14	Comoros	2000
	23	1007CIV	Ivory Coast	1999
	24	760GAB	Gabon	1998
	25	1049NGA	Nigeria	2000
	26	962UGA	Uganda	1998
	27	139aCOL	Colombia	1990
	28	CP2	Colombia	2009
	29	CRB10	Costa Rica	1999
	30	CUB A 6,1	Cuba	2009
	31	GUY1-2	French Guyana	
	32	449HND	Honduras	1993
	33	H-G -1	Honduras	2009
	34	616JAM	Jamaica	1995
	35	RD11	Dominican Rep.	1999
	36	101CHN	China	1988
	37	689FJI	Fidji	1996
	38	132IND	Indonesia	1990
	39	684IDN	Indonesia	1996
	40	574MYS	Malaysia	1995
	41	297PNG	Papua New Guinea	1989
	42	447PNG	Papua New Guinea	1988
	43	398PHL	Philippines	1993
	44	515VNM	Viet Nam	1994

$$\begin{cases} \frac{dS}{dt} = b - dS - \beta(I_+ + I_-)S - \varepsilon I_+ I_- S \\ \frac{dI_+}{dt} = \beta I_+ S + \frac{\varepsilon}{2} I_+ I_- S - (d + \alpha) I_+ \\ \frac{dI_-}{dt} = \beta I_- S + \frac{\varepsilon}{2} I_+ I_- S - (d + \alpha) I_- \end{cases} \quad (3)$$

To simplify the analysis we now define  $I$  as the total density of infected leaf surface units ( $I = I_+ + I_-$ ) and  $x$  the proportion of leaf surface units that are infected by pathogens with the first sexual sign, say  $+$  ( $x = I_+/I$ ).  $1 - x$  is therefore the proportion of leaf surface that is infected by pathogens with the other sexual sign ( $-$ ). The system becomes:

$$\begin{cases} \frac{dS}{dt} = b - dS - \beta IS - \varepsilon x(1 - x)I^2 S \\ \frac{dI}{dt} = \beta IS + \varepsilon x(1 - x)I^2 S - (d + \alpha)I \\ \frac{dx}{dt} = \varepsilon x(1 - x) \left( \frac{1}{2} - x \right) IS \end{cases} \quad (4)$$

Equilibrium densities of healthy and infected plant surface units can be found by equating the left-hand terms of system (4) to 0 and solving for  $S$ ,  $I$  and  $x$ . Stability analyses of these equilibriums allow determining criteria for disease spread and disease level, depending on the different parameters values.

### **Artificial inoculation experiments**

We carried out artificial inoculations of strains of the two species on banana leaf pieces maintained in survival conditions in Petri dishes. We used 25 *M. fijiensis* strains and 19 *M. musicola* strains, sampled worldwide from sensible banana cultivars (Table 1). Each strain, or isolate, is derived from a single ascospore isolated from an infected banana leave. Inoculations were performed by spraying conidia suspensions on leaf pieces as described later. For each strain, 6 leaf fragments were inoculated with the same spore suspension, in order to obtain 6 replicates per strain. Data describing lesion appearance, growth and sporulation, were collected during the development of symptoms by direct visuals marks of lesions on the one hand, and digital image analysing on the other hand.



### Artificial inoculation procedure

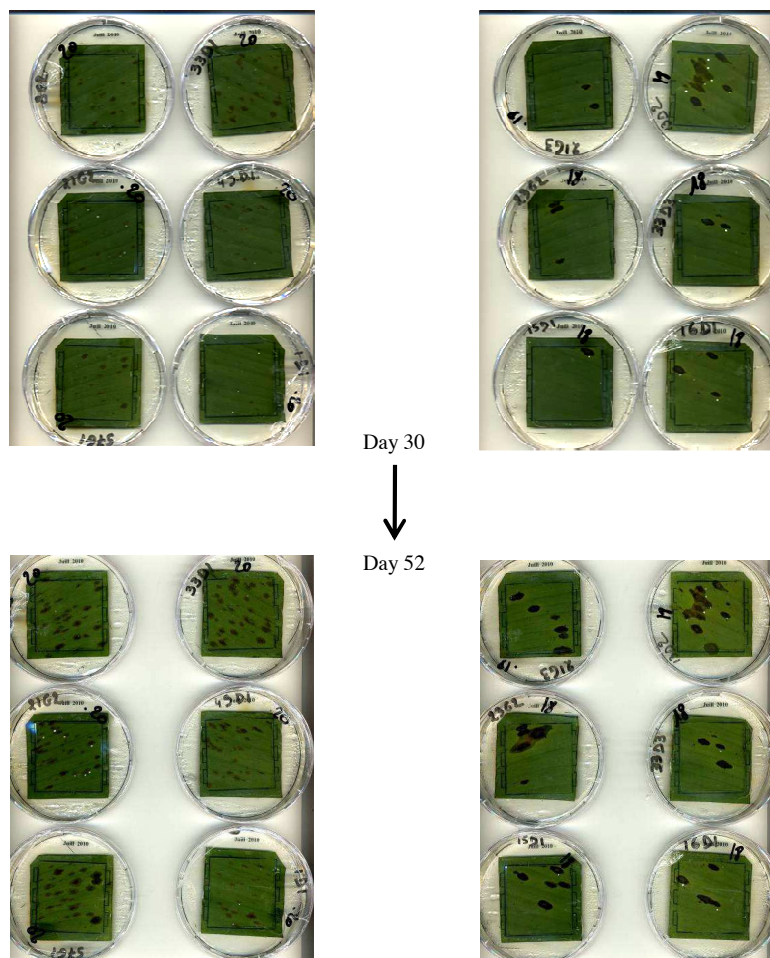
The banana leaf pieces inoculation procedure was adapted from Abadie *et al.* (2008), with a replacement of benzimidazole by gibberellic acid (5 mg/l) in the culture medium for surviving leaf pieces. For the preparation of surviving leaf pieces, around 30 sensible banana plants (Cavendish) derived from vitro-plants were cultivated in greenhouse for 6 to 7 months under optimal climatic and physiological conditions, before sampling of standardized fragments (25 cm<sup>2</sup>) from the first rank leaves (i.e., youngest deployed leaflets). Leaf pieces were placed in Petri dishes with the upper leaf surface down on a survival medium (Abadie *et al.* 2008). For the preparation of inoculum solutions, the 44 strains used were derived from a historical worldwide laboratory collection preserved in glycerol at -80°C. We followed the protocol steps of conidial inoculum preparation described in details in Abadie *et al.* (2008). The final step of conidia detachment from V8-sporulation medium cultures and collect in distilled water was improved here, by sonication of the distilled water (sonicator vibracell 72434, Bioblock), deposited on the colony surface. We verified with previous tests that sonication does not affect spore germination and mycelium growth (data not shown). Spore suspension concentrations were checked by visual counts in Malassez cells and adjusted to 5000 conidia/ml. Inoculations were performed by spraying 1mL of conidia suspension on leaf pieces in Petri dishes by using a microsyringe at constant pressure (Abadie *et al.* 2008). For each strain, inoculations were performed in 8 Petri dishes, following this order: previous control PDA dish without leaflet, then the six replicate dishes, then a final control PDA dish. All Petri dishes were incubated in a climatic chamber at 25°C with a 12-h light period. They were randomized and regularly mixed within climatic chamber to deal with undesirable potential environmental gradient. Germination rate was checked in the two control PDA dishes for each strain was controlled under binocular microscope after 2 days, with visual estimations of the density of germinating spores.

### Data collected

We followed up symptoms development in two different ways. During the first 4 weeks following symptoms onset (23 to 48 days after inoculation), all lesions were pointed with permanent colour markers every 3 days in average, in order to note the date of lesion appearance (visible to the naked eye) and the date of lesion conidiophores appearance (visible under microscope). After day 48 it was no longer possible to point lesions due to their surface

(a) *M. fijiensis*

(b) *M. musicola*



**Figure 3: Examples of Petri dishes with artificially inoculated leaf fragments, scanned at two different dates (day 30 and day 52).** The six replicates of two different strains are shown: (a) strain 20 for *M. fijiensis* and (b) strain 18 for *M. musicola*.

expansion. We thus obtained a dataset describing the time interval between lesions appearance and the onset of conidia production (corresponding to the difference between latency period and incubation period), and a dataset describing the increase of the total number of sporulating lesions in time. As not all strains have sporulated in laboratory conditions, these datasets were obtained from the 16 sporulating strains.

On the other hand, during 7 weeks (day 33 to day 77 after inoculation), Petri dishes were scanned every 3-4 days (Figure 3). Resulting images were analysed with IMAGEJ (<http://rsb.info.nih.gov/ij/>) by counting all lesions on leaf pieces at each date. For 8 strains (4 *M. fijiensis* and 4 *M. musicola*), the evolution of injured areas of 12 lesions per strain in time was also analysed with LAS (Leica Application Suite). These image analyses provided datasets describing the temporal dynamics of the total number of lesions, and evolution of lesions area.

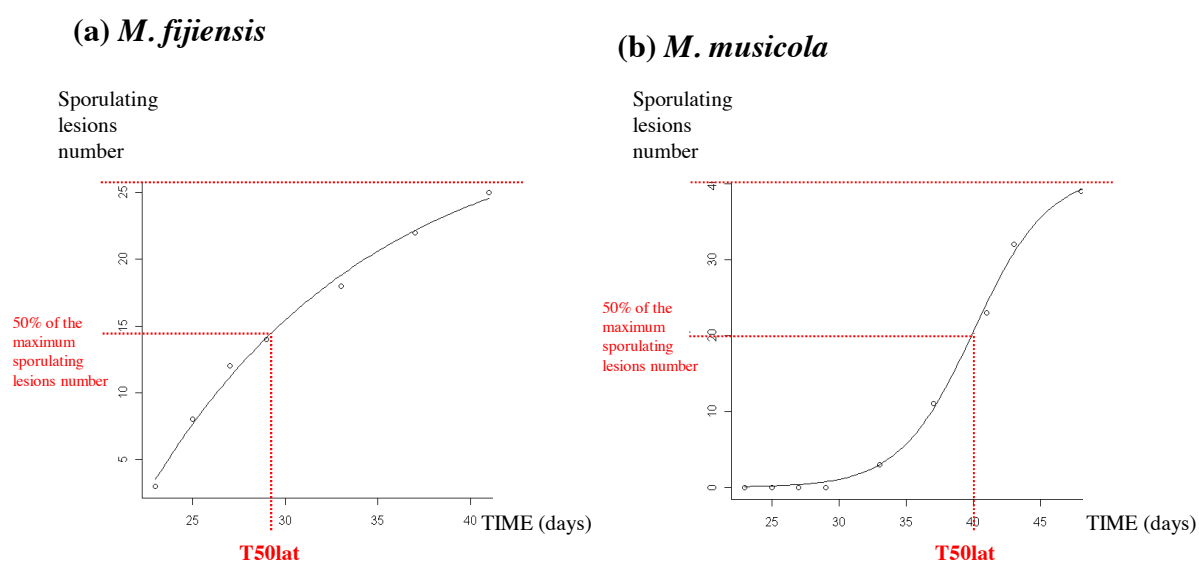
### Data analyses

All statistical analyses were performed with R (R Development Core Team 2011).

#### *Metrics for life history traits*

*Lesion growth dynamics.* Logistic growth curves were adjusted on data describing the increase of lesion number in the two species for every replicate of each strain. From these fits, we extracted a measure of incubation period using the time at which 50% of the total number of lesions have appeared (**T50inc**). We also extracted a proxy for infection efficiency using the maximal number of lesions (**LESmax**).

*Asexual sporulation dynamics.* The delay between lesion appearance and conidia sporulation (**DELAYspo**), obtained directly from the dataset of lesion colour marking, was used as a first proxy for latency. Besides, two different regression curves were adjusted for the two species on data describing the temporal dynamics of the proportion of sporulating lesions. Logistic growth curves were preferred for *M. musicola*, while for *M. fijiensis* lesions, monomolecular models were better adjusted (Figure 4). In both cases, a single curve was adjusted per strain on data pooled over the 6 replicates due to lack of power. From these fits, we extracted the time at which 50% of the final sporulating lesions have started producing conidia (**T50lat**) for each sporulating strain.



**Figure 4: Kinetics of total sporulating lesions number counted on the six leaf pieces for two different strains: (a) *M. fijiensis* strain 27 and (b) *M. musicola* strain 11.** Solid lines represent monomolecular and logistic estimated curves for *M. fijiensis* and *M. musicola* respectively. T50lat corresponds to the date at which 50% of the total number of sporulating lesions is achieved (represented here with schematic red lines).



### *Statistical tests on life history traits*

We firstly checked for undesirable associations between strain traits and strain sampling date using spearman correlation tests and any significant association was detected.

Differences between species were tested using linear mixed models, for **DELAYspo**, **T50inc**, and **LESmax**. As we disposed of several replicates per strain, we tested the significance of the fixed effect “species” and the random effect “strain” nested in “species”. To respect assumptions of normality and homoscedasticity of model residuals, we applied the log transformation on **DELAYspo** and **LESmax**. Differences between species were tested for **T50lat** using the non-parametric Mann-Whitney test. For lesions areas, we tested the differences between species by several Mann-Whitney tests performed for several dates of measure, on strain means. A global Wilcoxon signed rank test paired by date was also performed on data averaged per species.

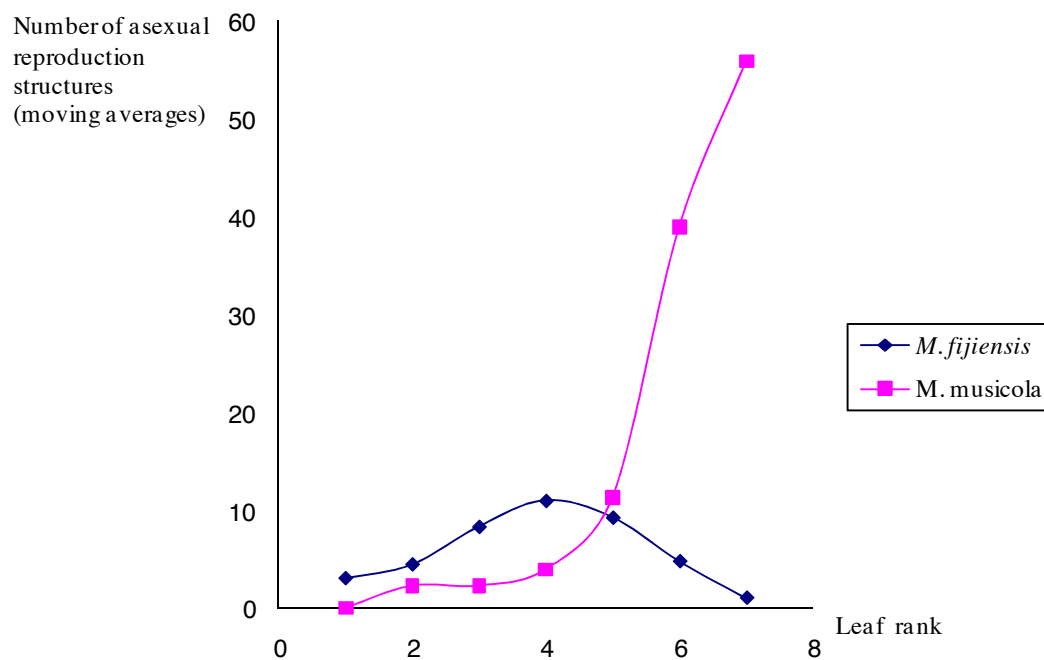
### **Natural infestation observations**

#### *Data collected*

Foliar samples were collected from natural infestation conditions in Martinique. Due to the very recent arrival of BLSD in Martinique, transitory cohabitation of the two diseases in the same plants could be observed there. A total of 10 banana plants in different smallholders with sufficient BLSD level were retained. Foliar pieces were collected from the left limb along all leaf ranks of bananas and treated then in laboratory. A total of 6 lesions of *M. fijiensis* and 6 lesions of *M. musicola* were cut up from all leaf samples, and bleached during a night in KOH (10%). *M. fijiensis* conidiophores and *M. musicola* sporodochia (i.e., dense fascicles of conidiophores) were counted under microscope. The data set therefore contained the number of asexual structures per lesion (i.e., conidiophores in *M. fijiensis* or sporodochia in *M. musicola*), for 6 lesions per leaf rank per species. Assuming that banana plants produce in average one leaf per week under tropical conditions, leaf rank can be considered as a good proxy for leaf age (in weeks).

#### *Species comparison*

Data describing number of asexual structures per lesion (i.e., conidiophores in *M. fijiensis* or sporodochia in *M. musicola*), were averaged over the 6 lesions collected per species and per



**Figure 5:** Graphical representation of the number of asexual structures per lesion (i.e., conidiophores in *M. fijiensis* or sporodochia in *M. musicola*, averaged over the 6 lesions collected per species and per leaf) according to leaves ranks of a single banana. Moving averages were computed over 3 successive values, resulting in smoothed curves.

leaf sample. The number of asexual structures varied too irregularly with leaf rank to properly fit growth curves. For each banana plant, we thus first computed a moving average of the number of conidiophores/lesion over 3 successive values in order to reduce stochastic noise (Figure 5). We then extracted the maximum number of conidiophores per lesion from the resulting smoothed curves and the rank of the corresponding leaf. For the following tests, the qualitative leaf rank factor was converted in quantitative time periods (number of weeks). Differences between species were tested for the two measures using the non-parametric Wilcoxon signed rank test paired by banana plant.

## RESULTS

### The model

#### Condition for disease spread

An important equilibrium solution is the “healthy host” equilibrium, corresponding to the loss of disease ( $I^* = 0$  and  $S^* = \frac{b}{d}$ ) explained in (2). For the disease to spread when rare, this equilibrium must be unstable. It can be shown that this condition is verified if:

$$\frac{b\beta}{d(d+\alpha)} > 1$$

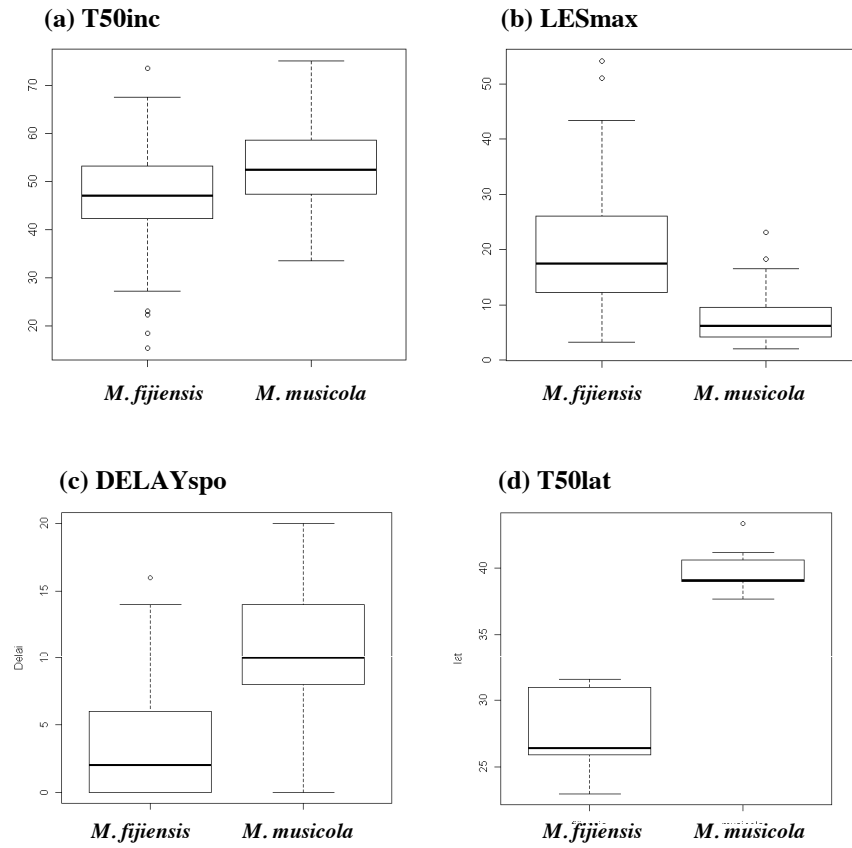
This condition defines the basic reproductive number:

$$R_0 = \frac{b\beta}{d(d+\alpha)},$$

determining whether the disease is able to spread or not. When  $R_0 < 1$ , the pathogen is eliminated and the host population grows to its carrying capacity. Conversely, if  $R_0 > 1$ , the pathogen population grows.  $R_0$  is negatively influenced by host basic mortality rate ( $d$ ) and disease virulence ( $\alpha$ ), and positively linked to host natality rate ( $b$ ) and asexual transmission rate ( $\beta$ ). Noticeably, the basic reproductive number  $R_0$  does not depend on sexual transmission rate ( $\varepsilon$ ) and is the same as expected under SI models without ascospore production.

#### System dynamics

Assuming that sexual spores are produced ( $\varepsilon \neq 0$ ), it can be shown that equilibriums with a loss of one or the other sexual sign ( $x = 0$  or  $x = 1$ ) are always unstable provided  $R_0 > 1$ . In



**Figure 6:** Graphical representations (boxplots) of distribution of values (averaged by strain) in each species, for the four life history traits measured in artificial inoculation experiments.  
 (a) T50inc, (b) LESmax, (c) DELAYspo, (d) T50lat

other words, a pathogen population that is able to spread on its host never loses sexual reproduction. However if one sexual sign is initially completely absent from the population, e.g., because it is lost at the introduction in a new environment, then the pathogen population spreads only by means of asexual spores at the condition that  $R_0 > 1$ .

When  $R_0 > 1$ , and provided that both sexual signs are initially present, host and pathogen populations may converge to an equilibrium given by:

$$x^* = \frac{1}{2}$$

$$I^* = \frac{1}{2} \left( \frac{b}{d + \alpha} - 4 \frac{\beta}{\varepsilon} + \sqrt{\left( \frac{b}{d + \alpha} + 4 \frac{\beta}{\varepsilon} \right)^2 - 16 \frac{d}{\varepsilon}} \right) \quad (6)$$

and

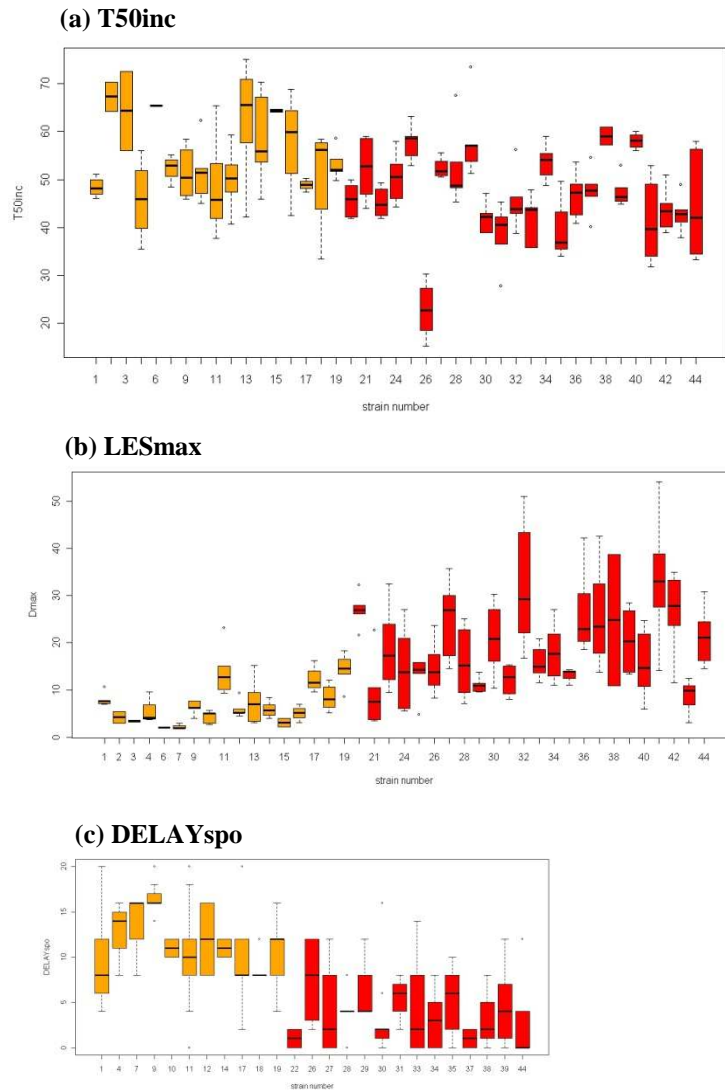
$$S^* = \frac{b - I^*(d + \alpha)}{d}. \quad (7)$$

From Equation 6, it is worth noticing that higher sexual transmission rate  $\varepsilon$  translates in higher infected plant surface ( $I^*$ ). The stability analysis of this equilibrium reveals that if the product  $bd$  is small, i.e., if the turn-over of healthy host surface units is low, and for intermediate sexual transmission rates  $\varepsilon$ , the equilibrium is unstable. In this latter situation the host and pathogen dynamics converge to a limit cycle.

### **Artificial inoculation experiments**

Differences between the two species explained a highly significant part of variance for all traits considered, except lesion area.

*M. fijiensis* was found to have a shorter incubation period (**T50inc** = 48 days for *M. fijiensis* and 58 days for *M. musicola*, p-value = 0.0019 for the species effect and p-value < 0.0001 for the strains effect in the linear mixed model), (Figure 6a). *M. fijiensis* tends to develop more lesions than *M. musicola* (**LESmax** = 17.9 lesions per leaf piece for *M. fijiensis* and 6.7 for *M. musicola*, p-value < 0.0001 for both species and strain effect in the linear mixed model), (Figure 6b). For lesion areas, *M. musicola* tended to display bigger lesions than *M. fijiensis*, although this effect was only detected with the global Wilcoxon signed rank test ( $V = 0$ ,



**Figure S1: Graphical representations (boxplots) of distribution of values in each strain, for the three life history traits for which we disposed of several replicated measures per strain : (a) T50inc, (b) LESmax, (c) DELAYspo**  
*M. musicola* strains (number 1 to 19) are represented in orange and *M. fijiensis* strains (number 20 to 44) are represented in red.

$p = 0.0078$ ) and not with the different Mann-Whitney tests performed for several dates of measure.

We also showed that *M. fijiensis* had a significantly shorter latency period than *M. musicola*, with also highly significant differences between strains. This was reflected by both a lower **DELAY<sub>spo</sub>** (3.7 days *M. fijiensis* and 11.2 days for *M. musicola*,  $p$ -value  $< 0.0001$  for both species and strain factors in the linear mixed model), (Figure 6c), and a lower **T50<sub>lat</sub>** (in average 28 days for *M. fijiensis* and 40 days for *M. musicola*, Mann-Whitney test:  $W=63$ ,  $p$ -value  $= 0.0010$ ), (Figure 6d).

### **Natural infestation observations**

We found that *M. fijiensis* total conidiophore production was in average 12 conidiophores per lesion, and mainly occurred on rank 5 leaves (after 5.1 weeks), whereas for *M. musicola*, a mean of 85 sporodochia per lesion were produced in average on rank 7-8 leaves. Assuming inoculations occurred on receptive cigar leaf for the two species, this represents a time interval of almost 3 weeks between *M. fijiensis* and *M. musicola* conidia production. Differences between species were highly significant for the two measures according to the Wilcoxon signed rank tests, ( $V = 0$ ,  $p = 0.00098$  for the maximum number of reproduction structures;  $V = 0$ ,  $p = 0.0068$  for the corresponding leaf rank).

## **DISCUSSION**

### ***The important role of asexual reproduction for colonizing new areas***

The model proposed here was formulated to better understand the contributions of conidia (asexual spores) and ascospores (sexual spores) in the dynamics of diseases such as those generated by *M. fijiensis* and *M. musicola*. A key insight provided by this model is that the basic reproductive number  $R_0$  (*i.e.*, the number of secondary lesions that an initial lesion will produce when the pathogen is rare) does not depend on ascospore production, but is only governed by asexual part of the life cycle as in the classical SI models (Gilligan & van den Bosch 2008). This theoretical result reflects the fact that sexual reproduction requires sufficient lesion densities for compatible lesions to coalesce and produce ascospores. When a pathogen is rare in a new area, the probability of encounter of two lesions of compatible





sexual sign is so small that only conidia production can have a substantial immediate role for pathogen invasion in such self-incompatible species.

Here the contribution of asexual spores was modeled through a transmission rate  $\beta$ , the rate at which healthy plant surface units become infected by asexual spores. This parameter  $\beta$  results from the combination of different life history traits. Asexual transmission rate is positively linked to asexual spore production capacities, spore probability of landing on receptive leaves, and infection efficiency of these spores. It is also negatively influenced by latency duration. An increase of  $R_0$  with asexual transmission rate means that, at low density, the pathogen will theoretically multiply and spread more and more efficiently as spore production capacities and/or infection efficiency increase, while latency period decreases. These traits may thus have a more important influence on pathogen fitness than previously thought for the two species (Brun 1963, Stover 1980, Gauhl 1994, Fouré *et al.* 1984), especially in context of emergence in new areas.

#### ***Asexual life strategies of M. fijiensis and M. musicola differ***

Substantial differences have been shown between *M. fijiensis* and *M. musicola*, regarding the combination of asexual life-history traits. The most striking result that emerges from our laboratory experiments is the huge conidia sporulation capacities of *M. musicola*, compared to *M. fijiensis*. Consequently, at first sight, *M. fijiensis* superiority is not obvious. *M. musicola* develops less but slightly larger lesions, and produces 5 to 10 times more reproduction structures per lesion than *M. fijiensis*. Previous phytopathological observations showed that reproduction structures corresponded for *M. fijiensis* to single conidiophores carrying on average 3 to 4 conidia, whereas for *M. musicola* reproduction structures corresponded to sporodochia gathering a hundred conidiophores, each one carrying a single conidia. Thus these differences could in principle lead to 100 to 1000 times more conidia in *M. musicola* than in *M. fijiensis*. Consistently with previous phytopathological observations (Stover 1980), orders of magnitude of conidia production achieve the huge number of several ten thousands conidia per lesion in *M. musicola*, while only a few hundred conidia per lesion in *M. fijiensis*. Such impressive differences, here measured from field samples, have also been qualitatively observed in our artificial inoculations, although not quantified.

Other traits could nevertheless favor *M. fijiensis* over *M. musicola*. *M. fijiensis* has a shorter latency period, producing its conidia more than 10 days earlier. This was verified both in artificial separate inoculations and in field measures in competition. In addition, *M. fijiensis*



develops symptoms more rapidly than *M. musicola* and achieves higher densities of primary lesions (i.e., lesions generated by the initial inoculated spores) for the same quantity of spore suspension inoculated. This could reflect a better infection efficiency of *M. fijiensis*. Phytopathological observations have consistently described a more important development of epiphyll mycelium able to colonize several stomata and thereby leading to multiple primary lesions from the same initial spore in *M. fijiensis* (Stover 1980).

In summary, *M. fijiensis*, despite being the best competitor, has not the best performances for all traits measured. A germinating spore of *M. musicola* seems to spend more time for vegetative growth and resources exploitation in the same leaf site, before developing important capacities of conidia production, whereas a germinating spore of *M. fijiensis* displays rapid vegetative colonization of leaf sites and earlier, but much lower, conidia production. This observation suggests that the two species are likely to face trade-offs in resource allocation to vegetative growth on the one hand and to reproduction on the other hand (Iwasa & Levin 1995, Reznick *et al.* 2000, Andanson 2010). Such results contrast with the majority of cases of strain replacements described in other fungal plant pathosystems (Milus & Line 1980, Miller *et al.* 1998, Milus *et al.* 2008, Pariaud *et al.* 2009b), notably regarding the low asexual sporulation capacities of *M. fijiensis*. For instance Miller *et al.* (1998) found that the newly established Columbian isolates of *Phytophthora infestans* had higher lesion expansion rates, shorter latent periods but also higher sporulation capacities than previous isolates. Pariaud *et al.* (2009b) suggested that the dominance of a new wheat leaf rust pathotype P1 on Soisson cultivar was also explained by its better values of life history traits, referring to larger lesion size, shorter latent period, longer infectious period, and again higher sporulation capacities. A possible explanation underlying this apparent absence of trade-offs in other pathogens is that these differ in their ability to extract resource from host. Some strains may have access to greater resource levels and therefore have better trait values on all traits compared to other strains with a lower resource extraction ability (Reznick *et al.* 2000). As a corollary, observing negative associations between *M. fijiensis* and *M. musicola* could suggest that they do not differ much in their abilities to access to host foliar resources. This is possible, as the major banana cultivar (Cavendish) is known as very susceptible with no partial resistance. Further interpretations will require additional experimental work.

Finally our results show how the effects of life history traits on fitness appear difficult to disentangle in such fungal pathogens. This renders more complex the possibility of using a simple metric as a proxy for fitness in filamentous fungi, as proposed by Pringle & Taylor (2002). For example we could suggest here that the quantity of asexual spore produced may



not be a suitable indicator of fitness for *M. fijiensis* at low densities, although it is generally one of the most intuitive fitness measure for fungal pathogens. Pringle & Taylor (2002) indeed insist on the need for a good knowledge of correlations and trade-offs between traits in the considered fungal model, before selecting a fitness measure. We think this is not an easy task at the light of our study, and especially as the pathosystems considered here are far from having the most complex life histories among fungal pathogens of plants. Let cite for instance the case of rusts species *Melampsora* sp. producing five spore stages during their life cycle (Pei *et al.* 2005). Much progress has to be made on these issues, especially by modeling explicitly the contribution of different spores types to diseases dynamics, as we proposed here.

***Could the differences in asexual life history strategy contribute to M. fijiensis competitive advantage?***

At first sight the very low number of conidia produced by *M. fijiensis*, as compared to *M. musicola*, renders its competitive superiority really questioning. The shorter latency (no more than 2 times shorter) and the higher infection efficiency of *M. fijiensis* (no more than 3 times more lesions) combined together are far from being sufficient to re-equilibrate the order of magnitude of the asexual transmission rate  $\beta$  between the two species. Nevertheless *M. fijiensis* is supposed to arrive at low density and start infecting banana in areas where *M. musicola* populations are already established. Hence both asexual and sexual reproductions are likely to contribute to *M. musicola* fitness. Accounting for competition would thus be a crucial extension of this model, in order to derive a measure of *M. fijiensis*  $R_0$  within an established population of *M. musicola*. Besides, in a context of competition, the asexual latency period and infection efficiency of *M. fijiensis* could in fact be key traits for its relative fitness. Latency period has already been suggested to be a crucial life history trait (Pariaud *et al.* 2009a) and it is commonly suggested that a pathogen able to colonize and occupy more rapidly and efficiency the available sites is likely to overcome a slower pathogen (Pringle & Taylor 2002). This could be true here in the particular context of the present pathosystem, where host is continually emitting new receptive leaves generating free sites available for both diseases. Under such conditions, competitive exclusion of one pathogen by the other could critically depend on the timing of reproduction rather than on the amount of spores produced. In a situation assuming rare *M. fijiensis* lesions, newly arrived, developing among dense *M. musicola* lesions, then the advantage of *M. fijiensis* asexual strategy must probably be viewed



on several successive generations: the higher infection efficiency of *M. fijiensis* might favor the rapid colonization of more sites by primary lesions, to be the first to occupy space. Earlier conidia production could accentuate this process by producing also secondary lesions more rapidly, as conidia dispersed by rain splash re-infect primarily inferior leaves and adjacent plants. Gilchrist *et al.* (2006) have proposed a competition model formulated initially for asexual saprobic fungi. They have shown that competitive exclusion of a strain 1 by a strain 2 occurs when the strain 2 is able to reduce the density of free sites under the threshold required for strain 1 dynamic equilibrium. Similar processes might contribute to *M. fijiensis* superiority, which would underline the counter-intuitive important contribution of asexual life history for *M. fijiensis* fitness advantage.

To go further, *M. fijiensis* success could also be enhanced by sexual reproduction as commonly thought, once population densities are high enough to allow lesions to coalesce. Considering our results we can raise the hypothesis that higher infection efficiency and rapid conidia re-infestation could favor efficient sexual reproduction, and might thus have an indirect fitness advantage in *M. fijiensis*. Lesions in higher densities are indeed supposed to coalesce more easily and rapidly, optimizing ascospore production. Field observations have shown that sexual reproduction occurs up to two weeks earlier in *M. fijiensis*. In addition, *M. fijiensis* would also have better sexual spore production capacities with an ascospore production of 2 to 10 times higher, according to Stover (1980). Considering the continuous emission of healthy receptive young leaves for ascospores by the host, this could confer a fitness advantage to *M. fijiensis* spread at the detriment of *M. musicola*. At the light of such results, it is not an easy task to disentangle the different factors likely to account for the rapid replacement of *M. musicola*, highlighting the need for further experimental and theoretical exploration of correlations and trade-offs between life-history traits, and particularly potential relations of synergy and/or trade-offs between asexual and sexual reproduction in such species.

Finally, another level of model complexity appears to be crucially lacking for understanding *M. fijiensis* successful spread across a banana production area. Wind-dispersed ascospores and rain-dispersed conidia have very differential dispersal capacities, ascospores being the main sources of inoculum at long distances in both species (Jones 2000, Rieux 2011). Moreover, in addition to the widely grown Cavendish banana in large-scale commercial plantations, *M. fijiensis* has other sensible hosts, especially plantains that display resistance to *M. musicola* and are thus generally almost free of SD (Jones 2000). Plantains fields represent an important





part of banana plantations in some regions (10% of the global commercial banana plantations, Abadie, personal data) and plantain plots for local consumption can be widespread in Africa. They could thus play an important role for *M. fijiensis* establishment. If we assume that *M. fijiensis* firstly establishes on plantains, avoiding thus its competitor, plantains field may then constitute important sources of *M. fijiensis* ascospores. Ascospores might thus arrive in mass into adjacent dessert banana plantations affected by *M. musicola*, by contrast with the situation of few propagules generating rare lesions we previously assumed. This could confer an additional fitness advantage to *M. fijiensis* in agro-ecosystems contexts presenting mosaics of Cavendish and Plantains. To better disentangle the potential contribution of such a phenomenon for *M. fijiensis* global invasion success, heterogeneity among host populations should also be considered in an explicitly spatial model at the scale of large banana production areas.

### ***Concluding remarks: our study in the context of biological invasions***

In conclusion, life histories of fungal pathogens obviously do not follow classical schemes. Can we suggest from our study that *M. fijiensis* displays life history traits more typical of the competitive K-strategy than *M. musicola*, as it has been described in other organisms (Facon *et al.* 2006, 2008, Duyck *et al.* 2007). The parallel with life strategies described for other species successively invading the same areas is not easy to draw. For example the lower quantities of asexual spores of *M. fijiensis*, and its great abilities for foliar sites occupation would be typical of K life history traits. On the contrary, its higher sexual sporulation and earlier asexual and sexual reproduction, added to the smaller size of its ascospores (Jones 2000), are rather typical of r-strategy, although we have showed here that they may confer a competitive advantage to *M. fijiensis*. It would seem here that we are facing subtle balances between asexual and sexual life histories. Moreover it would maybe be crucial to consider that the particular environment of the banana host plant represents for the pathogen by continuously producing new leaves totally free of disease, providing healthy sites available for the two diseases. Thus it can be advantageous in such a system to be the best pioneer, in other words to be a good colonist. Although more studies are required to further interpret our results, we highlight here how studies on fungal pathogen could bring new perspective for our understanding of species successive replacement, in the larger context of biological invasions. We support here that fungi are very interesting biological models able to challenge some general hypotheses of evolutionary biology drawn for plant or animal models.



## REFERENCES

- Abadie C, Zapater M-F, Pignolet L, Carlier J, Mourichon X (2008) Artificial inoculation on plants and banana leaf pieces with *Mycosphaerella* spp., responsible for Sigatoka leaf spot diseases. *Fruits* **63**, 319-323.
- Acebo-Guerrero Y, Hernandez-Rodriguez A, Heydrich-Perez M, El Jaziri M, Hernandez-Lauzardo AN (2012) Management of black pod rot in cacao (*Theobroma cacao* L.): a review. *Fruits* **67**, 41-48.
- Andanson A. (2010) Evolution de l'agressivité des champignons phytopathogènes, couplage des approches théorique et empirique. PhD thesis.
- Anderson PK, Cunningham AA, Patel NG, *et al.* (2004) Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution* **19**, 535-544.
- Arias P, Dankers C, Liu P, Pilhauskus P. (2003) The World bBanana Economy 1985-2002. *Roma: Food and Agriculture Organization of the United Nations*.
- Bahri B, Kaltz O, Leconte M, de Vallavieille-Pope C, Enjalbert J (2009) Tracking costs of virulence in natural populations of the wheat pathogen, *Puccinia striiformis* f.sp.tritici. *Bmc Evolutionary Biology* **9**.
- Brasier CM (2001) Rapid evolution of introduced plant pathogens via interspecific hybridization. *Bioscience* **51**, 123-133.
- Burt PJA, Rutter J, Gonzales H (1997) Short-distance wind dispersal of the fungal pathogens causing Sigatoka diseases in banana and plantain. *Plant Pathology* **46**, 451-458.
- Churchill ACL (2011) *Mycosphaerella fijiensis*, the black leaf streak pathogen of banana: progress towards understanding pathogen biology and detection, disease development, and the challenges of control. *Molecular Plant Pathology* **12**, 307-328.
- Cooke LR, Schepers HTAM, Hermansen A, *et al.* (2011) Epidemiology and Integrated Control of Potato Late Blight in Europe. *Potato Research* **54**, 183-222.
- de Vallavieille-Pope C, Giosue S, Munk L, *et al.* (2000) Assessment of epidemiological parameters and their use in epidemiological and forecasting models of cereal airborne diseases. *Agronomie* **20**, 715-727.
- Desprez-Loustau ML, Robin C, Buee M, *et al.* (2007) The fungal dimension of biological invasions. *Trends in Ecology & Evolution* **22**, 472-480.
- Desprez-Loustau ML, Courtecuisse R, Robin C, Husson C, Moreau PA, Blancard D, Selosse M-A, Lung-Escarmant B, Piou D, Sache I (2010) Species diversity and drivers of spread of alien fungi (sensu lato) in Europe with a particular focus on France. *Biological Invasions*, **12**(1):157-172
- Duyck P-F, David P, Quilici S (2007) Can more K-selected species be better invaders? A case study of fruit flies in La Reunion. *Diversity and Distributions* **13**, 535-543.
- Estoup A, Guillemaud T (2010) Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology* **19**, 4113-4130.
- Facon B, Genton BJ, Shykoff J, *et al.* (2006) A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology & Evolution* **21**, 130-135.
- Facon B, Pointier JP, Jarne P, Sarda V, David P (2008) High genetic variance in life-history strategies within invasive populations by way of multiple introductions. *Current Biology* **18**, 363-367.

- Fouré E, Grisoni M, Zurfluh R (1984) Les cercosporioses du bananier et leurs traitements. Comportement des variétés. Etude de la sensibilité variétale des bananiers et plantains à *Mycosphaerella fijiensis* Morelet et de quelques caractéristiques biologiques de la maladie des raies noires au Gabon. *Fruits* **39**, 365-378.
- Fullerton RA, Olsen TL (1995) Pathogenic variability in *Mycosphaerella-fijiensis* Morelet, cause of black sigatoka in banana and plantain. *New Zealand Journal of Crop and Horticultural Science* **23**, 39-48.
- Gauhl F (1994) Epidemiology and ecology of black sigatoka (PhD Thesis), INIBAP, Montpellier
- Gilchrist MA, Sulsky DL, Pringle A (2006) Identifying fitness and optimal life-history strategies for an asexual filamentous fungus. *Evolution* **60**, 970-979.
- Gilligan CA (2008) Sustainable agriculture and plant diseases: an epidemiological perspective. *Philosophical Transactions of the Royal Society B-Biological Sciences* **363**, 741-759.
- Gilligan CA, van den Bosch F (2008) Epidemiological models for invasion and persistence of pathogens. In: *Annual Review of Phytopathology*, pp. 385-418.
- Goyeau H, Halkett F, Zapater MF, Carlier J, Lannou C (2007) Clonality and host selection in the wheat pathogenic fungus *Puccinia tritricina*. *Fungal Genetics and Biology* **44**, 474-483.
- Hovmoller MS, Sorensen CK, Walter S, Justesen AF (2011) Diversity of *Puccinia striiformis* on Cereals and Grasses. In: *Annual Review of Phytopathology, Vol 49* (ed. VanAlfen NKBGLJE), pp. 197-217.
- Hovmoller MS, Yahyaoui AH, Milus EA, Justesen AF (2008) Rapid global spread of two aggressive strains of a wheat rust fungus. *Molecular Ecology* **17**, 3818-3826.
- Hufbauer, R. A., Facon, B., Ravigné, V., Turgeon, J., Foucaud, J., Lee, C. E., Rey, O., & A. Estoup (2011). Anthropogenically-Induced Adaptation to Invade (AIAI): Contemporary adaptation to human-altered habitats within the native range can promote invasions. *Evolutionary Applications*.
- Iwasa Y, Levin SA (1995) The timing of life-history events. *Journal of Theoretical Biology* **172**, 33-42.
- Jacome LH, Schuh W (1993) Effect of temperature on growth and conidial production in vitro and comparison of infection and aggressiveness in vivo among isolates of *Mycosphaerella fijiensis* var. *difformis*. *Trop. Agric. (Trinidad)* **70**, 51-59.
- Jacome LH, Schuh W (1993) Effect of temperature on growth and conidial production in-vitro, and comparison of infection and aggressiveness in-vivo among isolates of *mycosphaerella-fijiensis* var *difformis*. *Tropical Agriculture* **70**, 51-59.
- Jones DR (2000) Introduction to banana, abaca´ and enset. In: Diseases of Banana, Abaca´ and Enset (ed. Jones DR), pp. 37-79. CABI Publishing, Wallingford, UK.
- Marin DH, Romero RA, Guzman M, Sutton TB (2003) Black sigatoka: An increasing threat to banana cultivation. *Plant Disease* **87**, 208-222.
- Meredith DS (1970) Major banana diseases : past and present status. *Review of Plant Pathology* **49**, 539-554.
- Miller JS, Johnson DA, Hamm PB (1998) Aggressiveness of isolates of *Phytophthora infestans* from the Columbia Basin of Washington and Oregon. *Phytopathology* **88**, 190-197.
- Milus EA, Kristensen K, Hovmoller MS (2008) Increased aggressiveness of *Puccinia striiformis* f. sp *tritici* at least partially explains recent stripe rust epidemics. *Phytopathology* **98**, S107-S107.
- Milus EA, Line RF (1980) Characterization of resistance to leaf rust in pacific northwest wheats.

- Phytopathology* **70**, 167-172.
- Mouliom Pefoura A, Mourichon X (1990) Développement de *Mycosphaerella musicola* (maladie de Sigatoka) et *M. fijiensis* (maladie des raies noires) sur les bananiers et plantains. Etude du cas particulier des productions d'altitude. *Fruits* **45**, 17-24.
- Pariaud B, Ravigne V, Halkett F, *et al.* (2009) Aggressiveness and its role in the adaptation of plant pathogens. *Plant Pathology* **58**, 409-424.
- Pariaud B, Robert C, Goyeau H, Lannou C (2009) Aggressiveness Components and Adaptation to a Host Cultivar in Wheat Leaf Rust. *Phytopathology* **99**, 869-878.
- Parnell M, Burt PJA, Wilson K (1998) The influence of exposure to ultraviolet radiation in simulated sunlight on ascospores causing Black Sigatoka disease of banana and plantain. *International Journal of Biometeorology*, **42**, 22-27.
- Pei MH, Bayon C, Ruiz C (2005) Phylogenetic relationships in some *Melampsora* rusts on Salicaceae assessed using rDNA sequence information. *Mycological Research* **109**, 401-409.
- Pringle A, Taylor JW (2002) The fitness of filamentous fungi. *Trends in Microbiology* **10**, 474-481.
- Reznick D, Nunney L, Tessier A (2000) Big houses, big cars, superfleas and the costs of reproduction. *Trends in Ecology & Evolution* **15**, 421-425.
- Rhodes PL (1964) A new banana disease in Fiji. *Commonwealth Phytopathological News*, **10**, 38-41.
- Rieux A. (2011) Etude des processus de dispersion et des flux géniques chez un champignon phytopathogène : le cas de *Mycosphaerella fijiensis* à l'échelle d'un bassin de production Camerounais. PhD thesis, Montpellier SupAgro
- Roff, D.A. (1992). The Evolution of Life Histories: Theory and Analysis. *Chapman and Hall, New York*
- Romero RA, Sutton TB (1997) Reaction of four *Musa* genotypes at three temperatures to isolates of *Mycosphaerella fijiensis* from different geographical regions. *Plant Disease* **81**, 1139-1142.
- Singh RP, Hodson DP, Huerta-Espino J, *et al.* (2011) The Emergence of Ug99 Races of the Stem Rust Fungus is a Threat to World Wheat Production. In: *Annual Review of Phytopathology*, Vol 49 (ed. VanAlfen NKBGLJE), pp. 465-481.
- Stearns, S.C. (1992). The Evolution of Life Histories. Oxford University Press. 248 p.
- Stover RH (1980) Sigatoka leaf spots of bananas and plantains. *Plant Disease* **64**, 750-756.
- Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology* **46**, 75-100.



# **DISCUSSION GENERALE**





## DISCUSSION GENERALE et PERSPECTIVES

Jusqu'à récemment encore, la maladie des raies noires du bananier faisait partie de ces exemples typiques de maladies fongiques émergentes récemment détectées et décrites quant aux symptômes observés sur les cultures hôtes, mais dont on ne connaissait rien de la dynamique populationnelle et des processus d'expansion, alors même que l'émergence avait rapidement pris une ampleur mondiale. Suite aux premières études de génétique des populations ayant décrypté la structure des populations de *Mycosphaerella fijiensis* de l'échelle d'une plante à celle d'un continent (Carlier *et al.* 1996, Hayden *et al.* 2003, Rivas *et al.* 2004), nous avons été en mesure de mener les travaux qui sont l'objet de cette thèse.

Nos travaux ont donc permis de mettre à jour, à échelle globale, les routes d'invasion, incluant la localisation des populations sources et les différents scénarios d'introduction sur les continents envahis (Chapitre 1). Nous avons pu confirmer l'existence d'un centre majeur de diversité englobant l'Asie du Sud-Est et suggérons fortement que cette zone correspond (ou au moins englobe) le centre d'origine de *M. fijiensis*. Nous avons ensuite pu mettre en évidence un contraste remarquable entre les histoires africaine et américaine (qui sont les deux continents où nous disposons du plus grand nombre d'échantillons). Alors qu'un petit nombre d'individus très proches génétiquement seraient à l'origine de l'invasion du continent africain, probablement au cours d'un seul événement d'introduction, le continent américain aurait été envahi par des populations résultant d'un mélange entre différentes sources, qui aurait eu lieu au Honduras en accord avec les données historiques. La reconstruction des populations sources n'a pas pu être faite de façon précise mais nous pouvons dire que l'introduction africaine trouverait sa source en Asie du Sud-Est (pourtour de la Mer de Chine méridionale) tandis que les sources américaines pourraient provenir à la fois d'Asie du Sud-Est (plutôt PNG) et, de façon très supportée, en Océanie.

A échelle continentale, nous avons pu préciser les processus de dissémination au sein des continents envahis, et plus particulièrement les caractéristiques démographiques des populations et l'intensité des fluctuations démographiques passées dues au processus d'invasion (Chapitre 2). Les populations envahissantes auraient subi des goulots d'étranglement drastiques. Quelques hypothèses peuvent être formulées quant à la



dissémination à la fois graduelle et stochastique du pathogène, qui pourraient avoir une base à la fois naturelle et anthropique.

Enfin nous avons apporté des éclairages nouveaux sur la stratégie de vie de *M. fijiensis*, notamment la combinaison de THV liés au cycle asexué. Une efficacité d'infection supérieure, combinée à une sporulation asexuée plus précoce, bien que moins importante, seraient susceptible de conférer des avantages à *M. fijiensis* pour s'établir au sein d'agrosystèmes dominés par *M. musicola* (Chapitre 3).

On se propose d'intégrer l'ensemble des avancées effectuées sur chacun de ces aspects, pour avancer dans la compréhension des points exposés ci-après. Cette discussion sera l'occasion de souligner les limites de nos travaux, des suites à leur donner et des perspectives qu'ils ouvrent. Nous nous sommes posé les questions suivantes:

- ❖ A la lumière de nos résultats, que peut-on dire des mécanismes écologiques et évolutifs qui sous-tendraient l'invasion mondiale de *M. fijiensis* ? Comment pourraient-ils se replacer dans le contexte des scénarios de Facon *et al.* (2006) exposés en introduction ?
- ❖ Quelles questions et perspectives académiques pour la biologie évolutive plus générale émergent de cette étude de cas ?
- ❖ La raison d'être première de la phytopathologie, dans laquelle est ancrée notre étude, reste la prévention et le contrôle des maladies. Bien que les motivations principales de nos travaux aient été plutôt d'ordre fondamental, que peuvent-ils apporter de façon directement appliquée pour les stratégies de prévention et de lutte contre la maladie ?

### **Mécanismes d'invasion chez *M. fijiensis* : quels fondements écologiques et évolutifs ?**

Nous avons montré que *M. fijiensis*, lors de son établissement, supplantait son compétiteur et que pour cela, certains traits d'histoire de vie liés au cycle asexué, notamment l'efficacité d'infection et la latence, pourraient avoir plus d'importance que ce que l'on soupçonnait pour son succès d'invasion. La question centrale qui se pose si l'on veut se replacer dans le cadre commun des scénarios de Facon *et al.* (2006) serait donc de savoir si ces traits ont évolué



permettant ce succès dans les aires envahies, ou non. A l'issue de nos travaux à échelle mondiale, nous ne sommes pas directement en mesure de répondre à cette question, mais les éléments nouveaux apportés par cette thèse permettent d'avancer des hypothèses.

Dans les éléments qui pourraient être en faveur de cette évolution adaptative se trouve tout d'abord le délai (« lag time ») décrit dans les données historiques entre l'établissement de la maladie au Honduras et l'émergence sud-américaine (Stover & Dickson 1976). Ce délai pourrait résulter d'un effet Allee uniquement (i.e., limitation du taux de croissance de la population ici due à la difficulté de trouver des partenaires sexuels compatibles en faible densité), mais pourrait tout aussi bien correspondre à un temps d'adaptation des populations, en particulier suite au mélange de populations qui est associé à l'introduction de la maladie au Honduras. Notre étude expérimentale de mesure de traits d'histoire de vie suggère d'ailleurs l'existence d'un certain potentiel évolutif sur certains traits mesurés (temps d'incubation, efficacité d'infection, latence), ceci en raison des différences entre souches très significatives au sein d'une même espèce. Ces différences entre souches auraient donc une base génétique et pourraient refléter des différences entre populations à travers le monde, même si cela reste à étudier. Il a déjà été montré chez certaines espèces de champignons que la latence notamment, est un trait héritable qui peut être ajusté suivant les niveaux de compétition interspécifique (Pariaud *et al.* 2009a). Une adaptation de ces traits à l'arrivée au Honduras pourrait ainsi être envisagée.

D'un autre côté, les résultats centraux de nos travaux semblent fournir des arguments plutôt contre l'idée d'une adaptation des traits clés pour le succès d'invasion. L'énorme perte de diversité neutre observée en Afrique par succession de goulots d'étranglements drastiques sur une population fondatrice déjà appauvrie en diversité lors de l'introduction unique ne semble pas pour autant avoir affecté le succès d'invasion de *M. fijiensis* dans ce continent. De plus, la reconstruction des scénarios mondiaux d'invasion ne semble pas révéler de scénario « tête de pont » apparent, puisque les populations sources d'Afrique d'une part et d'Amérique d'autre part ne se situeraient même pas dans les mêmes régions géographiques. Cela n'a pas empêché, encore une fois la réussite de *M. fijiensis* sur tous les continents. L'existence d'une adaptation sur les traits d'histoire de vie supposerait alors qu'elle ait eu lieu à plusieurs reprises indépendamment, lors de scénarios d'introductions très variés, ce qui est moins facile à envisager. Enfin, bien que notre manque de puissance statistique nous oblige à être prudent, nous ne montrons pas de tendance à des différences entre aire d'origine et aires d'introduction



sur les traits d'histoire de vie. Si cela s'avère vrai, cela suggérerait que la stratégie de vie de *M. fijiensis* est la même sur tous les continents. Ces éléments suggèrent donc que la stratégie de vie de *M. fijiensis* était préalablement adaptée à être meilleure que celle de *M. musicola*. Ce pourrait être particulièrement plausible en Afrique et suggérerait que le mélange de populations n'a en rien favorisé l'invasion s'il s'avère que c'est aussi le cas en Amérique.

D'importantes limitations notamment techniques non résolues ici font encore obstacle pour aller plus loin sur cette question.

Tout d'abord, parce que notre expérience visait à comparer deux espèces, et non pas les aires natives et envahie au sein de l'espèce *M. fijiensis*, elle ne permet pas de tester des différences dans les valeurs de traits d'histoire de vie entre des isolats de l'aire d'origine et des différents continents envahis. Sur les 19 souches étudiées (14 asiatiques, 15 africaines et 15 américaines), aucune tendance ne semble se dégager. Des études de génétique quantitative sur des populations seraient naturellement requises pour explorer cette question.

Il s'ajoute à cela qu'en raison de l'immensité de l'aire de distribution de *M. fijiensis* encore trop peu échantillonnée (notamment l'aire d'origine), nous n'étions pas en mesure d'établir de lien génétique précis entre les populations envahissantes et leurs populations sources. Ne pouvant donc comparer rigoureusement ce qui est vraiment comparable, nous sommes limités pour la mise en œuvre de méthodes de génétique quantitative comparant les traits d'histoire de vie par des approches de type  $Q_{st}/F_{st}$  (Keller & Taylor 2008).

Enfin, le dernier obstacle (mais peut-être le plus important, à résoudre avant tous les autres) est que l'on connaît encore très mal la répartition des deux espèces *M. fijiensis* et *M. musicola* dans leurs aires d'origines, ni même la répartition de leurs hôtes cultivés et sauvages. Nous n'avons pas d'élément sur l'éventuelle coexistence ou l'interaction qu'il peut y avoir entre les deux espèces au sein de l'aire native. Par exemple on peut se demander si l'exclusion de *M. musicola* par *M. fijiensis* n'a lieu qu'au sein des agrosystèmes industriels.

En fait cette dernière question serait particulièrement cruciale. Considérant tous les éléments que nous venons d'exposer, le scénario d'AIAI (« Anthropogenically-Induced Adaptation to Invade ») récemment proposé par Hufbauer *et al.* (2011) pourrait tout à fait être à la base du succès d'invasion de *M. fijiensis*. Ce scénario propose que de récentes adaptations aux caractéristiques des milieux anthropisés aient lieu au sein même de l'aire d'origine des espèces, facilitant ensuite le succès d'invasion. L'introduction de ces populations préalablement adaptées aux milieux anthropiques dans de nouvelles aires géographiques a en





effet de grandes chances de se faire vers des milieux anthropisés très semblables. Il pourrait tout particulièrement rendre compte de nombreux cas d'invasions de pathogènes en agriculture en raison des profondes modifications des milieux engendrées par l'homme au sein des agrosystèmes, et de l'uniformisation de ces modifications à échelle mondiale.

Dans notre étude de cas, le scénario d'AIAI pourrait être envisagé à deux échelles temporelles différentes. Ce mécanisme évolutif aurait pu avoir lieu premièrement lors de la domestication ancienne du bananier et la mise en place des premières cultures qui auraient pu s'effectuer, comme nous l'avons suggéré dans un centre d'origine plus localisé que le Sud-Est Asiatique, la Papouasie-Nouvelle-Guinée étant notamment un bon candidat (Chapitre I). Une implantation des bananiers domestiqués dans tout le Sud-Est Asiatique aurait alors facilité la propagation et l'établissement de l'agent pathogène dans cette région. A une échelle plus contemporaine, les modifications drastiques des pratiques agricoles au cours du siècle dernier auraient pu déclencher, ou non, le même mécanisme d'adaptation à ces environnements dans l'aire pacifique. Si l'on suppose que *M. musicola* a pu devenir (par changement adaptatif ou simplement environnemental) particulièrement performant dans les cultures industrielles dès le début du siècle, générant ainsi son épidémie mondiale, on pourrait imaginer que *M. fijiensis* ait dû s'adapter pour le supplanter et s'établir lui-même au sein de tous les agrosystèmes industriels. De tels cultures industrielles ayant parallèlement été implantées sur tous les continents, l'intensification des échanges et l'introduction du pathogène sur les autres continents aurait été ainsi facilitée par son adaptation préalable à dominer l'agrosystème. Pour le savoir, des études de génétique quantitative doivent apporter la preuve d'une adaptation au sein de l'aire native, entre les environnements initiaux et les environnements modifiés par l'Homme, ainsi que la preuve que l'on retrouve cette même adaptation chez les populations envahissantes. Encore une fois, nous ne sommes pas en mesure de le démontrer proprement à l'issue de notre étude, mais cette hypothèse souligne la nécessité d'approfondir notre connaissance du pathogène dans son aire d'origine.

### **L'étude des champignons envahissants : de nouveaux éclairages pour la biologie évolutive**

Le Chapitre III, au détour de ses réflexions sur les stratégies d'histoire de vie de *M. fijiensis* et *M. musicola*, a soulevé bon nombre d'interrogations et hypothèses quant à la place des cycles de vie complexes des champignons phytopathogènes au sein de la théorie d'évolution des traits d'histoire de vie. Nous aimerions montrer par exemple si des compromis entre croissance végétative, et différents traits liés à la reproduction asexuée et



sexuée existent bel et bien au sein des deux espèces *M. fijiensis* et *M. musicola*, ou encore si *M. fijiensis* aurait une meilleure capacité d'extraction des ressources de l'hôte qui pourrait lui permettre d'optimiser de concert les valeurs de plusieurs traits (Reznick *et al.* 2000), comme on l'observe souvent dans les cas de remplacement successifs de pathogènes en agriculture. En effet, si *M. fijiensis* présente une production de conidies plus précoce mais beaucoup plus faible que *M. musicola*, il semble optimiser par ailleurs tous les traits d'histoire de vie liés à la reproduction sexuée, alliant grandes quantités d'ascospores et courtes périodes de latence. Aussi la tâche n'est-elle pas simple en raison de ces différents modes de reproduction, combinés qui plus est à différents traits de dispersion qui subiront eux aussi des pressions de sélection de l'environnement. Notons d'ailleurs que la dispersion, est également un trait d'histoire de vie clé qui fait l'objet de nombreuses recherches récentes dans le contexte des invasions, notamment quant à son évolution au cours des processus d'expansion (Travis & Dytham 2002, Travis *et al.* 2010).

Finalement, comme nous l'avons souligné dans la discussion de l'étude du Chapitre III, relier les résultats obtenus sur ces modèles aux stratégies r et K qui restent adaptées pour décrire bon nombre de modèles biologiques, peut devenir un casse-tête. Mais cela suscite aussi l'intérêt de mieux étudier ces modèles sous un angle fondamental, en alliant modèles théoriques et mesures de traits expérimentales. Comme l'ont déjà fait remarquer Desprez-Loustau *et al.* (2007), nous illustrons ici à quel point les champignons parasites sont en mesure de mettre au défi les hypothèses fondamentales de biologie évolutive fondées sur des modèles plantes et animaux. Ces organismes sont susceptibles d'apporter des éclairages nouveaux et méritent d'être mieux intégrés dans l'élaboration de cadres théoriques communs dans ces disciplines, et ceci, que l'on parle de la biologie des invasions ou de la biologie évolutive en général.

### **Les recherches fondamentales et leurs outils pour optimiser la lutte contre la maladie**

Cette étude nous a permis d'explorer et tester un certain nombre d'hypothèses historiques et épidémiologiques quant à la dissémination de la maladie et son établissement. Nous sommes finalement allés à l'encontre d'un certain nombre d'entre elles, en ayant par exemple rejeté l'hypothèse historique d'invasion Africaine à partir du Gabon, ou encore revaloriser le rôle potentiel de la reproduction asexuée pour le succès compétitif de *M.*



*fijiensis*. Or cette meilleure connaissance de la maladie n'est pas sans intérêt appliqué immédiat.

Les études de génétique des populations nous ont permis de mettre en évidence une forte structure de populations, même à des échelles intra-continentales, résultant de goulots d'étranglement particulièrement drastiques. Il était déjà suggéré que les introductions d'un continent à l'autre requéraient l'intervention de l'homme plutôt que celle des spores seules (Carlier *et al.* 1996). Nous avons montré de plus que ces événements de fondation fréquents au sein des continents pourraient résulter de sauts stochastiques de la maladie, ne suivant pas toujours une progression géographique logique, comme en Afrique. Ces événements stochastiques se superposeraient en fait à une dispersion graduelle de la maladie à l'échelle de grands bassins de productions (Halkett *et al.* 2010), les deux mécanismes étant susceptibles de contribuer à l'invasion rapide des continents. Or, s'il est vrai que les données génétiques ne permettent pas de différencier événements naturels de dispersion à longue distance par les ascospores (voire mouvements de fragments de feuilles infectées dispersés par des événements climatiques) et transports de matériel végétal infecté par les mouvements humains, on peut cependant discuter de la probabilité de ces deux phénomènes. L'établissement de la maladie suite à un événement de dispersion naturel requiert l'arrivée d'au moins deux spores de signes compatibles (de façon plus réaliste d'un petit groupe de spores) sur un même plant. A échelle très locale, démêler la part des deux phénomènes semble effectivement compliqué puisque qu'on ne connaît pas encore de manière assez précise la distance jusqu'à laquelle les spores peuvent établir de nouvelles populations (Rieux 2011). Cependant, un patron complètement stochastique de dispersion à grande échelle comme semble nous le suggérer les analyses ABC en Afrique, nous apparaît résulter plus probablement des mouvements anthropiques que des dispersions de groupes d'ascospores. Un certain nombre de centres d'études destinées à l'étude et la gestion des maladies des bananiers se sont en effet implantés dans plusieurs pays africains dans les années 60-70s. Chargés de l'amélioration variétale du bananier, ceux-ci ont constitué des collections de plants de bananiers et pourraient bien avoir leur part de responsabilité dans les sauts de maladies à très grande distance. Ceci pourrait d'ailleurs expliquer pourquoi l'introduction initiale sur le continent africain semblerait avoir eu lieu directement au cœur du continent et non sur la côte. Si l'on ajoute au cas africain celui de l'introduction en Amérique, concordante avec les données historiques de création d'une collection d'envergure de *Musa* au Honduras (Rosales *et al.* 1999), et enfin le cas de l'introduction de la maladie en Martinique, il en ressort que



notre étude pointe fortement du doigt les responsabilités des activités et échanges anthropiques pour l'invasion mondiale aussi rapide de la maladie. Au premier abord ceci n'était pas forcément intuitif chez un pathogène dont les quantités colossales de spores produites peuvent être transportées par le vent à grande distance. L'idée est communément répandue que l'on ne pourra de toute façon pas lutter contre la propagation par le vent de ces maladies fongiques, compliquant la bonne application des mesures de quarantaine. Or nous appuyons ici l'importance de respecter ces mesures de prévention et de contrôle, puisque les capacités de dispersion naturelles du pathogène semblent moins importantes que prévu, comparées aux transports humains.

Il nous semble important d'insister sur le fait que ces éclairages n'auraient pas pu être apportés sans la mise en œuvre de l'ensemble des méthodes de génétique des populations, mais aussi de façon peut-être moins attendue, des réseaux phylogénétiques basés sur des séquences. Le signal non homoplasique des séquences, et apte à conserver plus longtemps la trace de polymorphisme partagé, s'est avéré très précieux ici pour corroborer les analyses populationnelles des marqueurs microsatellites. Fitzpatrick *et al.* (2012) font la critique, dans un article récent, de l'emploi de la génétique des populations pour examiner les questions propres aux expansions stochastiques et évolutions très rapides caractérisant les invasions biologiques, suggérant que ces méthodes ne peuvent que rarement être appliquées à ces situations de déséquilibres démographiques des populations, sur des échelles de temps aussi courtes, et n'ayant pas d'application pour la lutter concrètement contre ce phénomène. Il nous semble que les quelques éléments que nous venons de mettre en lumière chez *M. fijiensis* supportent déjà l'intérêt que peut avoir l'emploi de méthodes de génétique des populations, même sur les échelles de temps récentes des invasions de pathogènes de plante, alors même que notre étude n'avait pas prétention à utiliser ces analyses dans une perspective directe de contrôle de la maladie.

Un autre éclairage nouveau apporté par nos travaux concerne l'importance de la partie végétative et asexuée dans la stratégie de vie de *M. fijiensis*. Alors que la production de conidies chez *M. fijiensis* apparaissait négligeable que ce soit en comparaison de sa production ascosporée ou de la production conidienne de son compétiteur *M. musicola*, nous avons souligné l'importance primordiale de la reproduction asexuée chez *M. fijiensis*. Celle-ci apparaît non seulement indispensable pour l'établissement initial de la maladie dans une aire





encore vierge, mais elle pourrait l'être aussi dans la capacité à s'établir au sein de plants déjà infectés par la maladie de Sigatoka jusqu'à être en mesure de l'exclure complètement. Ainsi la capacité du mycélium à coloniser efficacement la feuille, ajouté à la production plus précoce des conidies (période de latence beaucoup plus réduite), multiplie la densité des lésions primaires et secondaires chez *M. fijiensis* et pourrait de façon synergique optimiser alors sa reproduction sexuée. Il ne serait pas dénué de sens d'imaginer que les traits favorables (efficacité d'infection et latence conidienne) liés à la partie asexuée du cycle de *M. fijiensis* puissent alors impacter favorablement l'ensemble du processus, de l'établissement et l'exclusion compétitive de *M. musicola* au sein de chaque plant de bananier, jusqu'à la progression spatiale de la maladie par un front graduel d'ascospores à l'échelle de toute une parcelle. A nouveau, une telle hypothèse était au départ loin d'être acceptée au vu de la production modeste des conidies de *M. fijiensis* vouées à ne ré-infester que les feuilles voisines sur quelques mètres (Marin *et al.* 2003).

Or cibler les traits d'histoire de vie du pathogène qui apparaissent clés pour sa valeur sélective dans l'environnement, et donc pour son succès d'invasion, peut s'avérer particulièrement bénéfique pour la sélection de variétés partiellement résistantes de ses plantes hôtes. En effet la création variétale doit s'appuyer sur des composantes de résistance qui contrecarrent le pathogène à des étapes clés de son cycle infectieux, autrement dit ici de son histoire de vie. Si l'on peut démontrer par exemple que la latence asexuée s'avère être un trait primordial pour le succès d'invasion de *M. fijiensis*, des composantes de résistance capable d'augmenter ce temps de latence pourraient se révéler particulièrement efficaces. Afin que le pathogène ne contourne pas rapidement les résistances mises au point, une bonne connaissance de l'architecture génétique et du potentiel évolutif de ces traits pourra également requise.

A nouveau, d'un point de vue méthodologique, notre étude supporte ici l'intérêt des analyses fondamentales ancrées dans des questions de biologie évolutive, pour être en mesure de développer des stratégies de contrôle optimales et durables. Ceci doit passer par l'emploi intégré des modèles théoriques et de l'expérimentation pour examiner ces questions chez des organismes à l'histoire de vie aussi complexe. Comprendre quels sont les traits clés qui peuvent sous-tendre les émergences successives de pathogènes de plus en plus agressifs dans des agrosystèmes modernes, requiert bien sûr de poursuivre dans la compréhension de ces processus chez *M. fijiensis*, mais aussi plus d'explorer et confronter plus d'études de cas.



Enfin il est intéressant de souligner que, si nous ne sommes pas en mesure d'étudier le rôle des mélanges de populations pour le succès d'invasion en Amérique à l'issue de cette étude, nous le serions peut-être à une autre échelle, celle des Caraïbes. La situation de contact secondaire entre deux fronts génétiquement différenciés qui se présente en Martinique devrait ouvrir d'intéressantes perspectives pour poser ces questions. La mesure de traits d'histoire de vie chez des populations de chacun des deux fronts, ainsi que chez des populations issues de mélange (si cela s'avère confirmé), pourrait permettre d'évaluer les risques d'évolution de l'agent pathogène via ces mélanges de populations. Si l'on venait à montrer, comme cela a déjà été le cas chez quelques autres organismes (par ex. Lavergne & Molofsky 2007, Facon *et al.* 2011), que le mélange de populations potentialise le succès d'invasion du pathogène et donc les ravages qui en découlent, ce serait un argument de plus pour maintenir les mesures de quarantaine même après une première introduction.

Finalement l'ensemble de ces travaux appuie totalement l'idée d'une phytopathologie multidisciplinaire alliant notamment agronomie, génétique des populations et épidémiologie chez les pathogènes de plantes, comme cela a été souligné par de nombreux auteurs (Milgroom & Peever 2003, Gilligan 2008, Archie *et al.* 2009).



## REFERENCES

- Abadie C, Zapater M-F, Pignolet L, Carlier J, Mourichon X (2008) Artificial inoculation on plants and banana leaf pieces with *Mycosphaerella* spp., responsible for Sigatoka leaf spot diseases. *Fruits* **63**, 319-323.
- Acebo-Guerrero Y, Hernandez-Rodriguez A, Heydrich-Perez M, El Jaziri M, Hernandez-Lauzardo AN (2012) Management of black pod rot in cacao (*Theobroma cacao* L.): a review. *Fruits* **67**, 41-48.
- Andanson A. (2010) Evolution de l'agressivité des champignons phytopathogènes, couplage des approches théorique et empirique. PhD thesis.
- Anderson PK, Cunningham AA, Patel NG, *et al.* (2004) Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution* **19**, 535-544.
- Archie EA, Luikart G, Ezenwa VO (2009) Infecting epidemiology with genetics: a new frontier in disease ecology. *Trends in Ecology & Evolution* **24**, 21-30.
- Arias P, Dankers C, Liu P, Pilhauskus P. (2003) The World bBanana Economy 1985-2002. *Roma: Food and Agriculture Organization of the United Nations*.
- Arnold ML (2004) Natural hybridization and the evolution of domesticated, pest and disease organisms. *Molecular Ecology* **13**, 997-1007.
- Arzanlou M, Groenewald JZ, Fullerton RA, *et al.* (2008) Multiple gene genealogies and phenotypic characters differentiate several novel species of *Mycosphaerella* and related anamorphs on banana. *Persoonia* **20**, 19-37.
- Avisé JC (2009) Phylogeography: retrospect and prospect. *Journal of Biogeography* **36**, 3-15.
- Bahri B, Kaltz O, Leconte M, de Vallavieille-Pope C, Enjalbert J (2009) Tracking costs of virulence in natural populations of the wheat pathogen, *Puccinia striiformis* f.sp.tritici. *Bmc Evolutionary Biology* **9**.
- Banke S, McDonald BA (2005) Migration patterns among global populations of the pathogenic fungus *Mycosphaerella graminicola*. *Molecular Ecology* **14**, 1881-1896.
- Banke S, Peschon A, McDonald BA (2004) Phylogenetic analysis of globally distributed *Mycosphaerella graminicola* populations based on three DNA sequence loci. *Fungal Genetics and Biology* **41**, 226-238.
- Barney JN, Whitlow TH (2008) A unifying framework for biological invasions: the state factor model. *Biological Invasions* **10**, 259-272.
- Barres B, Halkett F, Dutech C, *et al.* (2008) Genetic structure of the poplar rust fungus *Melampsora larici-populina*: Evidence for isolation by distance in Europe and recent founder effects overseas. *Infection Genetics and Evolution* **8**, 577-587.
- Barrès, B., Carlier, J., Fenouillet, C., Seguin, M., Cilas, C., & V. Ravigné. Understanding the recent colonization history of a plant pathogenic fungus using population genetic tools and Approximate Bayesian Computation. In revision in **Heredity**.
- Barrett LG, Thrall PH, Burdon JJ, Linde CC (2008) Life history determines genetic structure and evolutionary potential of host-parasite interactions. *Trends in Ecology & Evolution* **23**, 678-685.
- Beaumont MA (1999) Detecting population expansion and decline using microsatellites. *Genetics* **153**, 2013-2029.

- Beaumont MA, Zhang WY, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics* **162**, 2025-2035.
- Beheregaray LB (2008) Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology* **17**, 3754-3774.
- Bertorelle G, Benazzo A, Mona S (2010) ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular Ecology* **19**, 2609-2625.
- Biek R, Real LA (2010) The landscape genetics of infectious disease emergence and spread. *Molecular Ecology* **19**, 3515-3531.
- Birch PRJ, Whisson SC (2001) *Phytophthora infestans* enters the genomics era. *Molecular Plant Pathology* **2**, 257-263.
- Brasier CM (2001) Rapid evolution of introduced plant pathogens via interspecific hybridization. *Bioscience* **51**, 123-133.
- Brewer MT, Milgroom MG (2008) Phylogeography and sequence diversity of genetic lineages of the grapevine powdery mildew fungus, *Erysiphe* (*Uncinula*) *necator*, in North America, Europe, and Australia. *Phytopathology* **98**, S25-S26.
- Brewer MT, Milgroom MG (2010) Phylogeography and population structure of the grape powdery mildew fungus, *Erysiphe necator*, from diverse *Vitis* species. *Bmc Evolutionary Biology* **10**.
- Brito P, Edwards SV (2009) Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica* **135**, 439-455.
- Brito PH, Edwards SV (2009) Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica* **135**, 439-455.
- Brown JKM, Hovmoller MS (2002) Epidemiology - Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science* **297**, 537-541.
- Brumfield RT, Beerli P, Nickerson DA, Edwards SV (2003) The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecology & Evolution* **18**, 249-256.
- Burt PJA, Rutter J, Gonzales H (1997) Short-distance wind dispersal of the fungal pathogens causing Sigatoka diseases in banana and plantain. *Plant Pathology* **46**, 451-458.
- Burton OJ, Phillips BL, Travis JMJ (2010) Trade-offs and the evolution of life-histories during range expansion. *Ecology Letters* **13**, 1210-1220.
- Byers JE (2000) Competition between two estuarine snails: Implications for invasions of exotic species. *Ecology* **81**, 1225-1239.
- Caffier V, Didelot F, Pumo B, *et al.* (2010) Aggressiveness of eight *Venturia inaequalis* isolates virulent or avirulent to the major resistance gene Rvi6 on a non-Rvi6 apple cultivar. *Plant Pathology* **59**, 1072-1080.
- Carlier J, Lebrun MH, Zapater MF, Dubois C, Mourichon X (1996) Genetic structure of the global population of banana black leaf streak fungus, *Mycosphaerella fijiensis*. *Molecular Ecology* **5**, 499-510.
- Carlisle DJ, Cooke LR, Watson S, Brown AE (2002) Foliar aggressiveness of Northern Ireland isolates of *Phytophthora infestans* on detached leaflets of three potato cultivars. *Plant Pathology* **51**, 424-434.
- Chakraborty R, Jin L (1993) A unified approach to study hypervariable polymorphisms: statistical considerations of determining relatedness and population distances. *Exs* **67**, 153-175.

- Churchill ACL (2010) *Mycosphaerella fijiensis*, the black leaf streak pathogen of banana: progress towards understanding pathogen biology and detection, disease development, and the challenges of control. *Molecular Plant Pathology* **12**, 307-328.
- Cooke LR, Schepers HTAM, Hermansen A, *et al.* (2011) Epidemiology and Integrated Control of Potato Late Blight in Europe. *Potato Research* **54**, 183-222.
- Cornuet J-M, Ravigne V, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *Bmc Bioinformatics* **11**.
- Cornuet J-M, Santos F, Beaumont MA, *et al.* (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* **24**, 2713-2719.
- Couch BC, Fudal I, Lebrun MH, *et al.* (2005) Origins of host-specific populations of the blast pathogen *Magnaporthe oryzae* in crop domestication with subsequent expansion of pandemic clones on rice and weeds of rice. *Genetics* **170**, 613-630.
- Crawford KM, Whitney KD (2010) Population genetic diversity influences colonization success. *Molecular Ecology* **19**, 1253-1263.
- Csillery K, Blum MGB, Gaggiotti OE, Francois O (2010) Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology & Evolution* **25**, 410-418.
- Day JP, Shattock RC (1997) Aggressiveness and other factors relating to displacement of populations of *Phytophthora infestans* in England and Wales. *European Journal of Plant Pathology* **103**, 379-391.
- de Bellaire LdL, Foure E, Abadie C, Carlier J (2010) Black Leaf Streak Disease is challenging the banana industry. *Fruits* **65**, 327-342.
- de Meeus T, McCoy KD, Prugnolle F, *et al.* (2007) Population genetics and molecular epidemiology or how to "debusquer la bete". *Infection Genetics and Evolution* **7**, 308-332.
- de Vallavieille-Pope C, Giosue S, Munk L, *et al.* (2000) Assessment of epidemiological parameters and their use in epidemiological and forecasting models of cereal airborne diseases. *Agronomie* **20**, 715-727.
- Desprez-Loustau ML, Robin C, Buee M, *et al.* (2007) The fungal dimension of biological invasions. *Trends in Ecology & Evolution* **22**, 472-480
- Desprez-Loustau ML, Courtecuisse R, Robin C, Husson C, Moreau PA, Blancard D, Selosse M-A, Lung-Escarmant B, Piou D, Sache I (2010) Species diversity and drivers of spread of alien fungi (*sensu lato*) in Europe with a particular focus on France. *Biological Invasions*, 12(1):157-172
- Dlugosch KM, Hays CG (2008) Genotypes on the move: some things old and some things new shape the genetics of colonization during species invasions. *Molecular Ecology* **17**, 4583-4585.
- Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology* **17**, 431-449.
- Dutech C, Enjalbert J, Fournier E, *et al.* (2007) Challenges of microsatellite isolation in fungi. *Fungal Genetics and Biology* **44**, 933-949.
- Dutech C, Fabreguettes O, Capdevielle X, Robin C (2010) Multiple introductions of divergent genetic lineages in an invasive fungal pathogen, *Cryphonectria parasitica*, in France. *Heredity* **105**, 220-228.

- Dutech, C., Barrès, B., Bridier, J., Robin, C., Milgroom, M. & V. Ravigné. The Chestnut Blight Fungus World Tour: successive introduction events from diverse origins in an invasive plant fungal pathogen. In press in *Molecular Ecology*.
- Duyck P-F, David P, Quilici S (2007) Can more K-selected species be better invaders? A case study of fruit flies in La Reunion. *Diversity and Distributions* **13**, 535-543.
- Eales J, Thorpe RS, Malhotra A (2010) Colonization history and genetic diversity: adaptive potential in early stage invasions. *Molecular Ecology* **19**, 2858-2869.
- Engelbrecht CJB, Harrington TC, Steimel J, Capretti P (2004) Genetic variation in eastern North American and putatively introduced populations of *Ceratocystis fimbriata* f. *platani*. *Molecular Ecology* **13**, 2995-3005.
- Estoup A, Guillemaud T (2010) Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology* **19**, 4113-4130.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611-2620.
- Evans EW (2004) Habitat displacement of North American ladybirds by an introduced species. *Ecology* **85**, 637-647.
- Excoffier L, Heckel G (2006) Computer programs for population genetics data analysis: a survival guide. *Nature Reviews Genetics* **7**, 745-758.
- Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* **1**, 47-50.
- Facon B, Crespín L, Loiseau A, *et al.* (2011) Can things get worse when an invasive species hybridizes? The harlequin ladybird *Harmonia axyridis* in France as a case study. *Evolutionary Applications* **4**, 71-88.
- Facon B, Genton BJ, Shykoff J, *et al.* (2006) A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology & Evolution* **21**, 130-135.
- Facon B, Hufbauer RA, Tayeh A, *et al.* (2011) Inbreeding Depression Is Purged in the Invasive Insect *Harmonia axyridis*. *Current Biology* **21**, 424-427.
- Facon B, Pointier JP, Jarne P, Sarda V, David P (2008) High genetic variance in life-history strategies within invasive populations by way of multiple introductions. *Current Biology* **18**, 363-367.
- Fahleson J, Nakyanzi M, Okori P, *et al.* (2009) Genetic analysis of *Mycosphaerella fijiensis* in the Ugandan Lake Victoria region. *Plant Pathology* **58**, 888-897.
- Fitzpatrick BM, Fordyce JA, Niemiller ML, Reynolds RG (2012) What can DNA tell us about biological invasions? *Biological Invasions* **14**, 245-253.
- Fouré E (1985) Les cercosporioses du bananier et leurs traitements. Comportement des variétés. Etude de la sensibilité variétale des bananiers et plantains à *Mycosphaerella fijiensis* Morelet au Gabon (maladie des raies noires). (suite III). *Fruits* **40**, 393-399.
- Fouré E, Grisoni M, Zurfluh R (1984) Les cercosporioses du bananier et leurs traitements. Comportement des variétés. Etude de la sensibilité variétale des bananiers et plantains à *Mycosphaerella fijiensis* Morelet et de quelques caractéristiques biologiques de la maladie des raies noires au Gabon. *Fruits* **39**, 365-378.
- Frankham R (2005) Invasion biology - Resolving the genetic paradox in invasive species. *Heredity* **94**, 385-385.



- Frossard P (1980) Apparition d'une nouvelle et grave maladie foliaire des bananiers et plantains au Gabon : la maladie des raies noires : *Mycosphaerella fijiensis* Morelet. *Fruits* **35**, 519-527.
- Fullerton RA, Olsen TL (1995) Pathogenic variability in *Mycosphaerella-fijiensis* Morelet, cause of black sigatoka in banana and plantain. *New Zealand Journal of Crop and Horticultural Science* **23**, 39-48.
- Gauhl F (1994) Epidemiology and ecology of black sigatoka (PhD Thesis), INIBAP, Montpellier
- Garniergere P, Dillmann C (1992) A computer-program for testing pairwise linkage disequilibria in subdivided populations. *Journal of Heredity* **83**, 239-239.
- Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology* **10**, 305-318.
- Gilchrist MA, Sulsky DL, Pringle A (2006) Identifying fitness and optimal life-history strategies for an asexual filamentous fungus. *Evolution* **60**, 970-979.
- Gilligan CA (2008) Sustainable agriculture and plant diseases: an epidemiological perspective. *Philosophical Transactions of the Royal Society B-Biological Sciences* **363**, 741-759.
- Gilligan CA, van den Bosch F (2008) Epidemiological models for invasion and persistence of pathogens. In: *Annual Review of Phytopathology*, pp. 385-418.
- Gladieux P, Zhang XG, Afoufa-Bastien D, *et al.* (2008) On the Origin and Spread of the Scab Disease of Apple: Out of Central Asia. *Plos One* **3**.
- Gladieux P, Zhang X-G, Roldan-Ruiz I, *et al.* (2010) Evolution of the population structure of *Venturia inaequalis*, the apple scab fungus, associated with the domestication of its host. *Molecular Ecology* **19**, 658-674.
- Goldstein DB, Linares AR, Cavallisforza LL, Feldman MW (1995) An evaluation of genetic distances for use with microsatellite loci. *Genetics* **139**, 463-471.
- Gottwald TR, Hughes G, Graham JH, Sun X, Riley T (2001) The citrus canker epidemic in Florida: The scientific basis of regulatory eradication policy for an invasive species. *Phytopathology* **91**, 30-34.
- Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity* **86**, 485-486.
- Goyeau H, Halkett F, Zapater MF, Carlier J, Lannou C (2007) Clonality and host selection in the wheat pathogenic fungus *Puccinia tritricina*. *Fungal Genetics and Biology* **44**, 474-483.
- Gruenwald NJ, Goss EM (2011) Evolution and Population Genetics of Exotic and Re-Emerging Pathogens: Novel Tools and Approaches. In: *Annual Review of Phytopathology, Vol 49* (ed. VanAlfen NKBGLJE), pp. 249-267.
- Guillemaud T, Beaumont MA, Ciosi M, Cornuet JM, Estoup A (2010) Inferring introduction routes of invasive species using approximate Bayesian computation on microsatellite data. *Heredity* **104**, 88-99.
- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* **52**, 696-704.
- Gurevitch J, Fox GA, Wardle GM, Inderjit, Taub D (2011) Emergent insights from the synthesis of conceptual frameworks for biological invasions. *Ecology Letters* **14**, 407-418.
- Halkett F, Coste D, Platero GGR, *et al.* (2010) Genetic discontinuities and disequilibria in recently established populations of the plant pathogenic fungus *Mycosphaerella fijiensis*. *Molecular Ecology* **19**, 3909-3923.
- Handley LJL, Estoup A, Evans DM, *et al.* (2011) Ecological genetics of invasive alien species. *Biocontrol* **56**, 409-428.

- Hare MP (2001) Prospects for nuclear gene phylogeography. *Trends in Ecology & Evolution* **16**, 700-706.
- Hayden HL, Carlier J, Aitken EAB (2003) Genetic structure of *Mycosphaerella fijiensis* populations from Australia, Papua New Guinea and the Pacific Islands. *Plant Pathology* **52**, 703-712.
- Henry P, Le Lay G, Goudet J, *et al.* (2009) Reduced genetic diversity, increased isolation and multiple introductions of invasive giant hogweed in the western Swiss Alps. *Molecular Ecology* **18**, 2819-2831.
- Hey J, Machado CA (2003) The study of structured populations - New hope for a difficult and divided science. *Nature Reviews Genetics* **4**, 535-543.
- Hovmoller MS, Sorensen CK, Walter S, Justesen AF (2011) Diversity of *Puccinia striiformis* on Cereals and Grasses. In: *Annual Review of Phytopathology, Vol 49* (ed. VanAlfen NKBGLJE), pp. 197-217.
- Hovmoller MS, Yahyaoui AH, Milus EA, Justesen AF (2008) Rapid global spread of two aggressive strains of a wheat rust fungus. *Molecular Ecology* **17**, 3818-3826.
- Hufbauer, R. A., Facon, B., Ravigné, V., Turgeon, J., Foucaud, J., Lee, C. E., Rey, O., & A. Estoup (2011). Anthropogenically-Induced Adaptation to Invade (AIAI): Contemporary adaptation to human-altered habitats within the native range can promote invasions. *Evolutionary Applications*.
- Iwasa Y, Levin SA (1995) The timing of life-history events. *Journal of Theoretical Biology* **172**, 33-42.
- Jacome LH, Schuh W (1993) Effect of temperature on growth and conidial production in vitro and comparison of infection and aggressiveness in vivo among isolates of *Mycosphaerella fijiensis* var. *difformis*. *Trop. Agric. (Trinidad)* **70**, 51-59.
- Jacome LH, Schuh W (1993) Effect of temperature on growth and conidial production in-vitro, and comparison of infection and aggressiveness in-vivo among isolates of *Mycosphaerella-fijiensis* var *difformis*. *Tropical Agriculture* **70**, 51-59.
- Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403-1405.
- Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* **11**.
- Jones DR (2000) Introduction to banana, abaca and enset. In: *Diseases of Banana, Abaca and Enset* (ed. Jones DR), pp. 37-79. CABI Publishing, Wallingford, UK.
- Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes* **5**, 187-189.
- Keller SR, Taylor DR (2008) History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. *Ecology Letters* **11**, 852-866.
- Keller SR, Taylor DR (2010) Genomic admixture increases fitness during a biological invasion. *Journal of Evolutionary Biology* **23**, 1720-1731.
- Khush GS (2001) Green revolution: the way forward. *Nature Reviews Genetics* **2**, 815-822.
- Kolar CS, Lodge DM (2001) Predicting invaders - Response. *Trends in Ecology & Evolution* **16**, 546-546.
- Kolar CS, Lodge DM (2002) Ecological predictions and risk assessment for alien fishes in North America. *Science* **298**, 1233-1236.

- Kolbe JJ, Glor RE, Schettino LRG, *et al.* (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature* **431**, 177-181.
- Lachmuth S, Durka W, Schurr FM (2011) Differentiation of reproductive and competitive ability in the invaded range of *Senecio inaequidens*: the role of genetic Allee effects, adaptive and nonadaptive evolution. *New Phytologist* **192**, 529-541.
- Lambrinos JG (2004) How interactions between ecology and evolution influence contemporary invasion dynamics. *Ecology* **85**, 2061-2070.
- Lavergne S, Molofsky J (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 3883-3888.
- Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology & Evolution* **17**, 386-391.
- Lee CE, Gelembiuk GW (2008) Evolutionary origins of invasive populations. *Evolutionary Applications* **1**, 427-448.
- Legg JP (1999) Emergence, spread and strategies for controlling the pandemic of cassava mosaic virus disease in east and central Africa. *Crop Protection* **18**, 627-637.
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451-1452.
- Linde CC, Zala M, McDonald BA (2009) Molecular evidence for recent founder populations and human-mediated migration in the barley scald pathogen *Rhynchosporium secalis*. *Molecular Phylogenetics and Evolution* **51**, 454-464.
- Lodge DM (1993) Biological invasions - lessons for ecology. *Trends in Ecology & Evolution* **8**, 133-137.
- Lombaert E, Guillemaud T, Cornuet JM, *et al.* (2010) Bridgehead Effect in the Worldwide Invasion of the Biocontrol Harlequin Ladybird. *Plos One* **5**.
- Lombaert E, Guillemaud T, Thomas CE, *et al.* (2011) Inferring the origin of populations introduced from a genetically structured native range by approximate Bayesian computation: case study of the invasive ladybird *Harmonia axyridis*. *Molecular Ecology* **20**, 4654-4670.
- Lucek K, Roy D, Bezault E, Sivasundar A, Seehausen O (2010) Hybridization between distant lineages increases adaptive variation during a biological invasion: stickleback in Switzerland. *Molecular Ecology* **19**, 3995-4011.
- Lumbsch HT, Buchanan PK, May TW, Mueller GM (2008) Phylogeography and biogeography of fungi. *Mycological Research* **112**, 423-424.
- Marin DH, Romero RA, Guzman M, Sutton TB (2003) Black sigatoka: An increasing threat to banana cultivation. *Plant Disease* **87**, 208-222.
- Marthey S, Aguilera G, Rodolphe F, *et al.* (2008) FUNYBASE: a FUNgal phYlogenomic dataBASE. *Bmc Bioinformatics* **9**.
- Meredith DS (1970) Major banana diseases : past and present status. *Review of Plant Pathology* **49**, 539-554.
- Meredith DS, Lawrence JS (1969) Black leaf streak disease of bananas (*Mycosphaerella fijiensis*) : symptoms of disease in Hawaii, and notes on the conidial state of the causal fungus. *Trans. Br. mycol. Soc.* **52**, 459-476.
- Milgroom MG, Peever TL (2003) Population biology of plant pathogens - The synthesis of plant disease epidemiology and population genetics. *Plant Disease* **87**, 608-617.

- Milgroom MG, Sotirovski K, Spica D, *et al.* (2008) Clonal population structure of the chestnut blight fungus in expanding ranges in southeastern Europe. *Molecular Ecology* **17**, 4446-4458.
- Miller JS, Johnson DA, Hamm PB (1998) Aggressiveness of isolates of *Phytophthora infestans* from the Columbia Basin of Washington and Oregon. *Phytopathology* **88**, 190-197.
- Milus EA, Kristensen K, Hovmoller MS (2008) Increased aggressiveness of *Puccinia striiformis* f. sp. *tritici* at least partially explains recent stripe rust epidemics. *Phytopathology* **98**, S107-S107.
- Milus EA, Line RF (1980) Characterization of resistance to leaf rust in pacific northwest wheats. *Phytopathology* **70**, 167-172.
- Montarry J, Andrivon D, Glais I, *et al.* (2010) Microsatellite markers reveal two admixed genetic groups and an ongoing displacement within the French population of the invasive plant pathogen *Phytophthora infestans*. *Molecular Ecology* **19**, 1965-1977.
- Montarry J, Hamelin FM, Glais I, Corbi R, Andrivon D (2010) Fitness costs associated with unnecessary virulence factors and life history traits: evolutionary insights from the potato late blight pathogen *Phytophthora infestans*. *Bmc Evolutionary Biology* **10**.
- Morrison DA (2005) Networks in phylogenetic analysis: new tools for population biology. *International Journal for Parasitology* **35**, 567-582.
- Mouliom Pefoura A, Mourichon X (1990) Développement de *Mycosphaerella musicola* (maladie de Sigatoka) et *M. fijiensis* (maladie des raies noires) sur les bananiers et plantains. Etude du cas particulier des productions d'altitude. *Fruits* **45**, 17-24.
- Mouliom-Pefoura A, Lassoudière A, Foko J, Fontem DA (1996) Comparison of development of *Mycosphaerella fijiensis* and *Mycosphaerella musicola* on banana and plantain in the various ecological zones in Cameroon. *The American Phytopathological Society* **80**, 950-954.
- Mourichon X, Fullerton RA (1990) Geographical distribution of the two species *Mycosphaerella musicola* Leach (*Cercospora musae*) and *M. fijiensis* Morelet (*C. fijiensis*), respectively agents of Sigatoka disease and Black Leaf Streak disease in bananas and plantains. *Fruits* **45**, 213-218.
- Munkacsi AB, Stoxen S, May G (2008) *Ustilago maydis* populations tracked maize through domestication and cultivation in the Americas. *Proceedings of the Royal Society B-Biological Sciences* **275**, 1037-1046.
- Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583-590.
- Neu C, Kaemmer D, Kahl G, Fischer D, Weising K (1999) Polymorphic microsatellite markers for the banana pathogen *Mycosphaerella fijiensis*. *Molecular Ecology* **8**, 523-525.
- Novak SJ (2007) The role of evolution in the invasion process. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 3671-3672.
- O'Donnell K, Kistler HC, Tacke BK, Casper HH (2000) Gene genealogies reveal global phylogeographic structure and reproductive isolation among lineages of *Fusarium graminearum*, the fungus causing wheat scab. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 7905-7910.
- Olivieri I (2009) Alternative mechanisms of range expansion are associated with different changes of evolutionary potential. *Trends in Ecology & Evolution* **24**, 289-292.
- Pariaud B, Ravigne V, Halkett F, *et al.* (2009) Aggressiveness and its role in the adaptation of plant pathogens. *Plant Pathology* **58**, 409-424.

- Pariaud B, Robert C, Goyeau H, Lannou C (2009) Aggressiveness Components and Adaptation to a Host Cultivar in Wheat Leaf Rust. *Phytopathology* **99**, 869-878.
- Parnell M, Burt PJA, Wilson K (1998) The influence of exposure to ultraviolet radiation in simulated sunlight on ascospores causing Black Sigatoka disease of banana and plantain. *International Journal of Biometeorology* **42**, 22-27.
- Pascual M, Chapuis MP, Mestres F, *et al.* (2007) Introduction history of *Drosophila subobscura* in the New World: a microsatellite-based survey using ABC methods. *Molecular Ecology* **16**, 3069-3083.
- Peever TL, Ibanez A, Akimitsu K, Timmer LW (2002) Worldwide phylogeography of the citrus brown spot pathogen, *Alternaria alternata*. *Phytopathology* **92**, 794-802.
- Pei MH, Bayon C, Ruiz C (2005) Phylogenetic relationships in some *Melampsora* rusts on Salicaceae assessed using rDNA sequence information. *Mycological Research* **109**, 401-409.
- Perrier X, De Langhe E, Donohue M, *et al.* (2011) Multidisciplinary perspectives on banana (*Musa spp.*) domestication. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 11311-11318.
- Philibert A, Desprez-Loustau M-L, Fabre B, *et al.* (2011) Predicting invasion success of forest pathogenic fungi from species traits. *Journal of Applied Ecology* **48**, 1381-1390.
- Posada D (2008) jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution* **25**, 1253-1256.
- Posada D, Crandall KA (2001) Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology & Evolution* **16**, 37-45.
- Pringle A, Taylor JW (2002) The fitness of filamentous fungi. *Trends in Microbiology* **10**, 474-481.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.
- Pujol B. & J.R. Pannell (2008). Reduced responses to selection after species range expansion. *Science* 321:96
- Pysek P, Jarosik V, Hulme PE, *et al.* (2010) Disentangling the role of environmental and human pressures on biological invasions across Europe. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 12157-12162.
- Pysek P, Richardson DM, Pergl J, *et al.* (2008) Geographical and taxonomic biases in invasion ecology. *Trends in Ecology & Evolution* **23**, 237-244.
- R Development Core Team (2011) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Raboin LM, Selvi A, Oliveira KM, *et al.* (2007) Evidence for the dispersal of a unique lineage from Asia to America and Africa in the sugarcane fungal pathogen *Ustilago scitaminea*. *Fungal Genetics and Biology* **44**, 64-76.
- Raymond M, Rousset F (1995) Genepop (version-1.2) - population-genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.
- Reznick D, Bryant MJ, Bashey F (2002) r- and K-selection revisited: The role of population regulation in life-history evolution. *Ecology* **83**, 1509-1520.
- Reznick D, Nunney L, Tessier A (2000) Big houses, big cars, superfleas and the costs of reproduction. *Trends in Ecology & Evolution* **15**, 421-425.

- Reznick DN, Ghalambor CK (2001) The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica* **112**, 183-198.
- Rhodes PL (1964) A new banana disease in Fiji. *Commonwealth Phytopathological News*, 10, 38-41.
- Richardson DM, Pysek P (2006) Plant invasions: merging the concepts of species invasiveness and community invasibility. *Progress in Physical Geography* **30**, 409-431.
- Richardson DM, Rejmanek M (2004) Conifers as invasive aliens: a global survey and predictive framework. *Diversity and Distributions* **10**, 321-331.
- Rieux A. (2011) Etude des processus de dispersion et des flux géniques chez un champignon phytopathogène : le cas de *Mycosphaerella fijiensis* à l'échelle d'un bassin de production Camerounais. PhD thesis, Montpellier SupAgro
- Rieux A, Halkett F, de Bellaire LD, *et al.* (2011) Inferences on pathogenic fungus population structures from microsatellite data: new insights from spatial genetics approaches. *Molecular Ecology* **20**, 1661-1674.
- Rivas GG, Zapater MF, Abadie C, Carlier J (2004) Founder effects and stochastic dispersal at the continental scale of the fungal pathogen of bananas *Mycosphaerella fijiensis*. *Molecular Ecology* **13**, 471-482.
- Robert S, Rieux A, Argout X, Carlier J, Zapater MF (2010) Optimized genotyping with microsatellite markers in the fungal banana pathogen *mycosphaerella fijiensis* (*mycosphaerellaceae*). *American Journal of Botany* **97**, E130-E132.
- Roff, D.A. (1992). *The Evolution of Life Histories: Theory and Analysis*. *Chapman and Hall, New York*
- Rosales FE, Tripon SC, Cerna J (eds) (1999) *A Tribute to the Work of Paul H. Allen: A Catalogue of Wild and Cultivated Bananas*. INIBAP/CIID/EARTH, Montpellier, France.
- Romero RA, Sutton TB (1997) Reaction of four *Musa* genotypes at three temperatures to isolates of *Mycosphaerella fijiensis* from different geographical regions. *Plant Disease* **81**, 1139-1142.
- Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* **4**, 137-138.
- Rosenthal DM, Ramakrishnan AP, Cruzan MB (2008) Evidence for multiple sources of invasion and intraspecific hybridization in *Brachypodium sylvaticum* (Hudson) Beauv. in North America. *Molecular Ecology* **17**, 4657-4669.
- Rousset F (2008) GENEPOP '007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**, 103-106.
- Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**, 365-386.
- Sakai AK, Allendorf FW, Holt JS, *et al.* (2001) The population biology of invasive species. *Annual Review of Ecology and Systematics* **32**, 305-332.
- Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9**, 615-629.
- Shea K, Chesson P (2002) Community ecology theory as a framework for biological invasions. *Trends in Ecology & Evolution* **17**, 170-176.

- Singh RP, Hodson DP, Huerta-Espino J, *et al.* (2011) The Emergence of Ug99 Races of the Stem Rust Fungus is a Threat to World Wheat Production. In: *Annual Review of Phytopathology, Vol 49* (ed. VanAlfen NKBGLJE), pp. 465-481.
- Stearns, S.C. (1992). The Evolution of Life Histories. Oxford University Press. 248 p.
- Storey JD (2002) A direct approach to false discovery rates. *Journal of the Royal Statistical Society Series B-Statistical Methodology* **64**, 479-498.
- Storz, J.F. and Beaumont, M.A. (2002)  
Testing for genetic evidence of population expansion and contraction: an empirical analysis of microsatellite DNA variation using a hierarchical Bayesian model. *Evolution*. **56**: 154-166.
- Stover RH (1978) Distribution and probable origin of *Mycosphaerella fijiensis* in southeast Asia. *Trop. Agric. (Trinidad)* **55**, 65-68.
- Stover RH (1980) Sigatoka leaf spots of bananas and plantains. *Plant Disease* **64**, 750-756.
- Stukenbrock EH, Banke S, Javan-Nikkhah M, McDonald BA (2007) Origin and domestication of the fungal wheat pathogen *Mycosphaerella graminicola* via sympatric speciation. *Molecular Biology and Evolution* **24**, 398-411.
- Stukenbrock EH, Banke S, McDonald BA (2006) Global migration patterns in the fungal wheat pathogen *Phaeosphaeria nodorum*. *Molecular Ecology* **15**, 2895-2904.
- Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology* **46**, 75-100.
- Sunnucks P (2000) Efficient genetic markers for population biology. *Trends in Ecology & Evolution* **15**, 199-203.
- Travis MJM, Dytham C (2002) Dispersal evolution during invasions. *Evolutionary Ecology Research* **4**, 1119-1129.
- Travis MJM, Muenkemueller T, Burton OJ (2010) Mutation surfing and the evolution of dispersal during range expansions. *Journal of Evolutionary Biology* **23**, 2656-2667.
- van Kleunen M, Dawson W, Schlaepfer D, Jeschke JM, Fischer M (2010) Are invaders different? A conceptual framework of comparative approaches for assessing determinants of invasiveness. *Ecology Letters* **13**, 947-958.
- Vila M, Weiner J (2004) Are invasive plant species better competitors than native plant species? evidence from pair-wise experiments. *Oikos* **105**, 229-238.
- Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.
- Whitney KD, Gabler CA (2008) Rapid evolution in introduced species, 'invasive traits' and recipient communities: challenges for predicting invasive potential. *Diversity and Distributions* **14**, 569-580.
- Wolfe LM, Blair AC, Penna BM (2007) Does intraspecific hybridization contribute to the evolution of invasiveness?: an experimental test. *Biological Invasions* **9**, 515-521.
- Yang B, James A, Zhong S (2007) Genetic diversity of *Mycosphaerella fijiensis* from bananas (*Musa* spp.) in Hawaii. *Phytopathology* **97**, S126-S127.
- Zaffarano PL, McDonald BA, Linde CC (2008) Rapid speciation following recent host shifts in the plant pathogenic fungus *Rhynchosporium*. *Evolution* **62**, 1418-1436.
- Zaffarano PL, McDonald BA, Linde CC (2009) Phylogeographical analyses reveal global migration patterns of the barley scald pathogen *Rhynchosporium secalis*. *Molecular Ecology* **18**, 279-293.

- Zapater MF, Duchemin M, Dussart JF, *et al.* (2008) Microsatellite markers for the fungal banana pathogens *Mycosphaerella fijiensis*, *Mycosphaerella musicola* and *Mycosphaerella eumusae*. *Molecular Ecology Resources* **8**, 1121-1125.
- Zhan J, McDonald BA (2004) The interaction among evolutionary forces in the pathogenic fungus *Mycosphaerella graminicola*. *Fungal Genetics and Biology* **41**, 590-599.
- Zhang DX, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology* **12**, 563-584.
- Zhou X, Burgess TI, De Beer ZW, *et al.* (2007) High intercontinental migration rates and population admixture in the sapstain fungus *Ophiostoma ips*. *Molecular Ecology* **16**, 89-99.